PSYCHOLOGICAL ROOM
GENERAL LIBRARY
UNIV. OF MICH.

FEB 16 1932

VOL. XXIII FEBRUARY 1932 NO. 2

The Journal of Educational Psychology

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching

HAROLD RUGG, *Editor*

CONTENTS

An Experimental Study of Certain Phases of a True-false Test . . 81
WILLIAM C. F. KRUEGER

A Comparison of Marks Based upon Weighted and Unweighted Items in a New-type Examination 92
EDWARD F. POTTHOFF AND NEVILLE E. BARNETT

Items of Gesell's Developmental Schedule Scaled 99
MINNIE L. STECKEL

Eye Movements and the Eye-hand Span in Typewriting 104
RUSSELL L. C. BUTSCH

A Comparison of the Results Obtained from the Teaching of Shorthand by the Word Unit Method and the Sentence Unit Method . . 122
MILDRED CLARK AND D. A. WORCESTER

The Retention of Certain Latin Syntactical Principles by First and Second Year Latin Students after Various Time Intervals . . . 132
LEO R. KENNEDY

The Bogey of the "Personal Equation" in Ratings of Intelligence . 147
HERBERT S. CONRAD

Book Reviews 150

Publications Received 159

$5.00 per Year · Published Monthly Except June to August by

WARWICK & YORK, INC.

BALTIMORE, MD.

Entered as Second Class Matter Nov. 15, 1921, at the Post Office at Baltimore, Md. under the Act of March 3, 1879; additional entry as second class matter at York, Pa.

Important New Texts

THE WHOLESOME PERSONALITY

A Contribution to Mental Hygiene

By WILLIAM H. BURNHAM, PH.D.

Professor Emeritus of Education and School Hygiene, Clark University; author of "The Normal Mind."

Dr. Burnham's new book on the development of personality is concerned chiefly with the course of normal personality and the conditions that favor its development. It discusses among other subjects, the sacredness of the individual's task, the scientific attitude, the hygiene of the social group. The emphasis is on positive methods. $2.75

THE MIND IN ACTION

A Study of Motives and Values

By A. CAMPBELL GARNETT

This third volume in the important new *Contemporary Library of Psychology*, deals with the problem of the sources of human wants and the determination of behavior and presents a new theory of volition, of habit, and of sentiment. $2.00

These are Appleton Books

D. APPLETON AND COMPANY

35 West 32nd Street New York

TEACHERS' HEADQUARTERS

European

Plan

Entirely

Fire-Proof

Centrally

Located

EDWARD DAVIS

Manager

HOTEL RENNERT

BALTIMORE, MARYLAND

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY

ESTABLISHED 1910

Editor: HAROLD RUGG

Board of Associate Editors

J. CARLETON BELL ARTHUR I. GATES RUDOLF PINTNER E. L. THORNDIKE

FRANK N. FREEMAN V. A. C. HENMON L. M. TERMAN GOODWIN B. WATSON

The price of the Journal is $5.00 a year in the United States; $5.30 in foreign countries. Subscriptions for less than a year are charged at the rate of 70 cents for each number ordered. Back volumes are $6.00 each; back issues, 90 cents each except when three or more years old, and then $1.00 each.

Subscribers should notify the publishers of change in address at least four weeks before publication of the issue with which the change is to take effect. Claims for non-receipt of an issue will not be honored unless made within two weeks after receipt of the next succeeding number.

Manuscripts for publication, books or other materials for review, and correspondence should be addressed to the publishers.

WARWICK AND YORK, INC., *Baltimore, Maryland*

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XXIII February, 1932 Number 2

AN EXPERIMENTAL STUDY OF CERTAIN PHASES OF A TRUE-FALSE TEST

WILLIAM. C. F. KRUEGER

College of the City of Detroit, Detroit, Michigan

This experiment was planned to suggest answers to the following problems:

I. Will a person write "true" more often than "false" when he guesses his answers at random in a true-false test? What proportion of these guesses are incidentally guessed correctly?

II. When a person is limited to fifty per cent of the guesses as "true" and fifty per cent of the guesses as "false," what proportion of the guesses will happen to be correct?

III. (*a*) What is the relation of degree of learning to the total frequency of the items marked correctly as "true" or "false"?

(*b*) What is the proportion of answers marked correctly as "true" and "false" when the subject is presented with different materials of varying degrees of difficulty? Are the proportional increases the same for three kinds of materials?

(*c*) What proportion of the answers marked as "true" or "false" will be marked correctly when the subject asserts that he is very certain that his answers are correct?

(*d*) To what extent does degree of learning, or knowledge of subject-matter, influence the frequency of actual correctness of judgment when the subject is very sure that he marks correctly?

PART I

Each member of a group of one hundred three undergraduates was presented with a set of lists of one hundred items to a list. The respective truth or falsity of every item was totally unknown to all members of the group. Each member, however, was asked to guess

an answer to each item as either "true" or "false." He was cautioned not to make his answers according to any scheme, for example, to alternate "true" and "false" while marking, or to check a certain number of items as "true" and then a certain number of items as "false" and so on until the list was completed. The subject was to refrain from counting his respective "true" and "false" answers. He simply was to write "true" or "false" after each item as whim and fancy suggested. The actual number of true and false statements was not made known to the subject. Was the average frequency of "true" fifty? Was it more, or less than fifty? Furthermore, how many of the items guessed as "true" and "false" coincided with a previously prepared key of fifty true and fifty false items?

Each subject received a list of one hundred words, a second list of one hundred nonsense syllables, and another list of one hundred three-place numbers. The words were monosyllabic nouns of three or four letters. The nonsense syllables were constructed in the ordinary manner, so that each syllable consisted of three different letters, a vowel between two consonants. The numbers were chosen so that no sequence, such as 123, or duplication as 424, occurred. The group was informed that later on the experimenter would read to them a series of words. Nonsense syllables, or numbers respectively, would be presented in like manner. These items to be read would be selected from the lists of one hundred before them. The subjects did not know which items had been selected or omitted in the making of the list to be read later. However, they were to select or guess which of the items the experimenter had selected. All words, syllables, or numbers, respectively, which the subjects thought would be selected, were to be marked by writing "true" after the items, and all items which they thought would not be included in this list to be read later, were to be marked as "false." Lists of nouns, syllables, and numbers had been selected because matter-of-fact statements might influence, by suggestion, an answer as "true" or "false." It was hoped that all suggesting influences would be avoided by selecting items in random order rather than by judging affirmatively or negatively worded statements. After each decision throughout the entire list, the subject covered up his writing to prevent his seeing what he had written before. Thus each subject made his one hundred guesses for each of the three lists. The results are shown in Table I.

In order to note how many of these guessed answers were guessed correctly as "true" and "false," the answers were checked by means

of a prepared key which contained the fifty selected items marked as true. The results are included in Table I.

TABLE I.—AVERAGE NUMBER OF ITEMS GUESSED AS "TRUE" AND AVERAGE NUMBER OF ITEMS GUESSED CORRECTLY "TRUE" OR "FALSE"

	AVERAGE	SD	PE
Number of items guessed as "true":			
Words	51.10	7.64	.50
Syllables	50.90	7.20	.48
Numbers	51.00	8.24	.54
Number of items guessed correctly:			
Words	49.87	6.34	.42
Syllables	50.02	6.42	.43
Numbers	49.75	6.14	.41

Whenever the subjects guessed "true" and "false" at random, and when they did not know the exact number of true items in the list, they wrote "true" after slightly more than fifty per cent of the items. The frequency of correctly guessed items, when checked by a key of fifty true and fifty false items, averaged almost fifty for the three lists. Although we found that the subjects wrote "true" more frequently than "false" under the conditions of the above experiment, we did not have the great difference as found in a similar investigation.[1] This disagreement may be attributed to the difference in the nature of the materials used. It is possible that the matter-of-fact statements suggested "true" more frequently than "false."

PART II

This experiment differed from the preceding arrangements in that the subjects were directed to limit their guesses of "true" to fifty and their guesses of "false" were to be fifty. The same one hundred three subjects were given a second set of the same lists of one hundred words, one hundred nonsense syllables, and one hundred three-place numbers. (It may be well to state here that the reading of the selected items, as promised in Part I, did not occur.) Thus the subjects still were forced to guess since they had no information upon

[1] Fritz (*Journal of Educational Psychology* Vol. XVIII, 1927, pp. 558–561) observed that students wrote "true" more frequently than "false" when guessing the answers. He found a ratio of 62 per cent to 38 per cent. "This tendency to give 'true' reactions is about the same whether we take actual class materials or whether we use experimental materials with which the students are totally unacquainted."

which to base their decisions. As before no particular scheme for making guesses was to be followed except that the limit of "true" and "false" guesses was to be fifty each. Again the answers were checked with the key in order to note the frequency and average distribution of correct guesses. The results are summarized in Table II.

TABLE II.—AVERAGE FREQUENCY OF CORRECTLY GUESSED ITEMS

Material	Average	SD	PE
Words	50.27	6.00	.40
Nonsense syllables	50.03	5.40	.36
Numbers	50.23	5.30	.35

The frequency of "true" was always fifty for all three lists in accord with the given directions. The average frequency of correctly guessed items was approximately fifty for the three types of materials. However, the SD's were consistently smaller than those for the same lists in Part I. This indicates that the range of frequencies was correspondingly nearer the average than that of Part I, but the average frequencies for the three lists were almost the same. If the scoring method of "right minus wrong" answers was used, the average scores would be approximately zero for all series of both Part I and Part II.

PART III

The original group of one hundred three subjects was divided into three sub-groups of approximately equal size and apparently equal learning ability. Each sub-group judged four lists of the one hundred words, one hundred nonsense syllables, and one hundred three-place numbers, using the same materials as before. One sub-group started with the words, the other with the syllables, and the third with the numbers, thus equalizing practice effects by rotation of the order in which the various materials were to be learned. At any one sitting the subjects were provided with four sheets of words, syllables, or numbers. The lists were placed on the desks, with the printed sides down. The subjects were informed that a list of fifty words, or syllables, or numbers, as required, would be read to them. They were to attempt to remember as many of these fifty items as possible, so that they could judge whether or not the items read were included in the list of one hundred before them. The fifty items were read at the

rate of two seconds per item. The words were not spelled, but the syllables were spelled by letter; and the three-place numbers were read by digits; for example, 387 was read as three-eight-seven. After the first reading one sheet was turned and the subjects were instructed to mark as follows.

The experimenter started to read each of the one hundred items in the order as printed on the sheet, at the rate of three seconds per item. Thus the subject was given three seconds for making a judgment and writing his answer. If he believed that the item on the list was one of the fifty items just read, he was to write "true" after the item. If he thought that the item was new or not among the fifty read, he was to write "false." If he did not know, he was asked to "guess" and write either "true" or "false." One condition of this experiment was that the judgment or guess was to be made within the three-second interval allowed for each item. No one was to "work ahead" not to leave blank spaces to be filled in later. If the subject was very certain that his answer was correct, the statement of "true" or "false" was to be underlined. This would mark especially those judgments of which the person was very sure. After completing the first sheet, these papers were collected. Then the subjects heard the same fifty items read as before, at the same rate but in a different order. A seond sheet was filled out according to the same instructions as for the first sheet. Again the subjects indicated, as before, whenever they were very certain of the correctness of judgment. After the completion of the second sheet, a third reading of the original fifty items was given, at the same rate but again in a different order. The respective items on the third sheet were filled in as before, keeping the three-second interval for judgment as the author read the one hundred items. As soon as these papers had been gathered, the fourth reading occurred. The last sheet was marked in accordance with the rules governing the previous procedures.

Each person worked on four lists of words, or four lists of syllables, or four lists of numbers on any one day. He took part in all of the three series. Before starting with Part III, each subject was given some preliminary practice by using lists of thirty items instead of one hundred. Otherwise the procedure was identical as outlined above, with the exception that the reading list contained only fifteen words, syllables, or numbers, respectively. The averages in the following calculations are based upon one hundred three scores for each average.

In Table III are found the average retention scores for the four degrees of learning based upon four frequencies of presenting the three kinds of materials. The data show that by increasing the frequency of readings, the respective average scores became larger. The successive average increases over the guess (taken at fifty) were 21.37, 10.77, 6.34, and 2.56 for the words lists after one, two, three, or four readings of the list of the fifty selected items. For the nonsense syllables, the successive increases were 12.14, 6.11, 5.62, and 3.86, while the corresponding scores for the numbers increased in order by 6.28, 2.25, 1.14, and .49 items respectively for the four frequencies of presentations.

The continuous increase in average retention score for each kind of material was negatively accelerated. The greatest absolute increases were with the words, and the smallest for the numbers lists. The list of words consisted of material already well known to the subjects, and no special learning was needed except for purposes of recognition. The nonsense syllables were almost entirely unknown to the group and thus more difficult to recognize. The syllables, constructed from the letters of the alphabet, had a considerable degree of similarity and therefore produced some confusion, and a possible inhibition effect. The three-place numbers were combinations

TABLE III.—AVERAGE NUMBER OF ITEMS SCORED CORRECTLY

	After one reading	After two readings	After three readings	After four readings
Words				
Average	71.64	82.41	88.75	91.31
SD	7.20	10.05	7.75	7.35
PE	.48	.67	.52	.49
Syllables				
Average	62.17	68.28	73.90	77.76
SD	6.55	7.70	8.60	9.70
PE	.44	.51	.57	.64
Numbers				
Average	56.51	57.76	58.90	59.39
SD	6.20	5.65	5.85	6.15
PE	.41	.38	.39	.41

selected from the digits 1 to 9 with the consequent high degree of similarity among the resultant numbers. This condition would tend to increase the possibilities of confusion and inhibition. The explanations offered may account for the three rates of increase for the three kinds of materials.

The subjects had been requested to underscore those answers of "true" and "false," concerning which, they were very sure that their judgments were correct. The average frequencies for the three materials and the four readings are shown in Table IV. After the first reading of the words, the average number of answers, underscored and correctly marked as "true" or "false," was 25.88, while that for the syllables and the numbers was 11.49 and 5.24, respectively. Since the meaningful words were better known and more readily recognized than either the nonsense syllables or the similar number combinations, the proportion of correctly marked answers was in relation to what the subject knew or remembered. This statement seems justified when the corresponding average retention scores in Table III are considered.

As the readings of the lists were repeated, thereby offering the subjects opportunities for additional learning, the average scores for

TABLE IV.—AVERAGE NUMBER OF ITEMS SCORED CORRECTLY AS "TRUE" AND AS "FALSE" WHEN THE SUBJECTS WERE VERY SURE OF THEIR JUDGMENTS

	After one reading	After two readings	After three readings	After four readings
Words				
Average	25.88	42.64	57.95	67.62
SD	14.40	19.10	22.85	22.15
PE	.96	1.27	1.52	1.47
Syllables				
Average	11.49	21.85	31.18	38.84
SD	9.90	14.90	17.20	20.30
PE	.66	.99	1.14	1.35
Numbers				
Average	5.24	10.25	12.91	16.12
SD	6.00	9.55	10.75	11.40
PE	.41	.63	.71	.76

the underlined answers became larger. After the fourth reading the average for the words was 67.62, for the syllables 38.84, and for the numbers 16.12. This indicated again that the more the subjects knew or had learned, the greater was the frequency of correct answers when the subjects felt very sure that they were correct in the judgments. The rate of increase tends towards a negative acceleration. The respective SD's for the four degrees of learning became larger as the number of readings increased. A study of the individual scores as compared to the class standing of the respective person showed rather outstandingly that the good students underlined fewer answers than the poor students. Furthermore, the good students made very few errors in judgment while the poor students had a very low percentage of correct scores. This difference became more marked as the readings continued. Both types of students underscored more answers as the learning process went on, but the rate of increase was much faster for the good students than for the poor students. Even after the fourth the percentage of errors for the good students was very small while that of the poor students was proportionately very much larger. Practically every large score of errors came from students who ranked rather low in class standing.

TABLE V.—AVERAGE NUMBER OF ITEMS SCORED INCORRECTLY AS "TRUE" AND AS "FALSE" WHEN THE SUBJECTS WERE VERY SURE OF THEIR JUDGMENTS

	After one reading	After two readings	After three readings	After four readings
		Words		
Average	3.45	3.46	3.56	3.49
SD	4.45	3.05	3.60	3.55
PE	.29	.20	.24	.23
		Syllables		
Average	2.63	3.86	4.51	4.58
SD	3.90	4.65	4.40	4.30
PE	.26	.30	.29	.29
		Numbers		
Average	2.17	4.17	5.37	6.88
SD	3.90	6.70	6.75	7.65
PE	.26	.45	.45	.67

It seemed also of interest to tabulate the frequency of underlined, but incorrect judgments, although the subjects believed that they were correct. We find the averages for the various conditions in Table V. For the words, the error of judgment was rather constant, ranging from averages of 3.45 to 3.56 for the frequencies of readings. The SD's were smaller for the later tests than for the first test. One might conclude, therefore, that when the material was well known and meaningful, the error of judgment was somewhat constant. This tendency did not hold true for the judgments of the syllables and the numbers. Here the frequency of error increased with the increasing frequency of readings and tests. When one notes the relatively small number of correct answers (see Table III) and the relatively small number of underscored answers (see Table IV) for the syllable and the number series, it is readily seen that the frequency of error could not be very large when the total frequency of "sure" judgments was small. This was especially obvious with the number series. As the readings of the lists (the fifty selected items read by the experimenter) were repeated, the subjects became more confident of their judgments, as indicated in Table IV. The rapid increase in error may be explained, in part at least, by the relative difficulty of the more readily confused materials. The more familiar the subjects were with the material to be learned, and the greater the similarity of the items among which to discriminate, the more likely an error of judgment would occur due to confusion and inhibition effects. When the subjects had mastered the meaning of the items, as in the case of the simple nouns, the proportional error was much smaller and more constant.[1]

---

[1] In order to test the factor of meaning, on a later occasion the writer continued the experiment with six of the original subjects. Four more alternate readings and markings were made. The frequency of correct answers continued to increase as the readings were repeated and the number of underscored answers became larger with each new test. The proportion of incorrectly marked and underlined items remained somewhat constant for the nouns, but decreased for the syllables after the first additional reading. The decrease came after the third additional reading of the numbers. The introspective reports of these subjects (although not trained in making observations for such reports) indicated that the "meaning" of an item was an essential basis for certainty of judgment. As soon as the syllables or number became associated with some extraneous but definite meaning, it was much easier to ascertain one's certainty of judgment. The data was not included here on account of the small number of subjects used. Furthermore, these six subjects happened to be students whose ranking was higher than the average of the group.

The fast rate of increase in error with the syllables and the numbers, in spite of repeated readings of the list of the fifty selected items, may be ascribed in part to the fact that the poor students made very large error scores. It seemed that in the cases of several subjects who ranked very low in class standing, increased number of readings caused greater confusion. This increase in error was much greater with the numbers than with the syllables. It appeared that the good students were more conservative and cautious in that they underlined fewer answers than the poor group and underlined fewer incorrect answers, particularly for the first test given.

Conclusions

The data suggest the following conclusions:

1. When the subjects guessed "true" or "false" at random, the tendency was to write "true" slightly more often than "false." The average frequency of items guessed correctly by chance was fifty per cent of the total number of guesses.

2. When the subjects were instructed to mark in random order 50 per cent of the guesses as "true" and 50 per cent of the guesses as "false," the average frequency of items guessed correctly by chance was fifty per cent of the total number of guesses.

3. As the number of presentations of materials (for purposes of learning) was increased, the average total frequencies of the items checked correctly as "true" or as "false," increased at a negatively accelerated rate.

4. The proportion of correctly marked items was dependent upon the relative difficulty of the materials. The more familiar the material was, the greater was the corresponding increase in retention score for the various frequencies of presenting the learning material.

5. When the subjects believed that their judgments were correct, the proportion of such answers (indicated by underlining the response word) was dependent upon the degree to which the subjects knew the meaning of the material to be learned, or had ascribed extraneous meaning to such material.

6. For meaningful, familiar material, such as the simple nouns, the number of errors was rather constant, irrespective of the number of times the material had been presented for learning.

7. The more meaningless the material, and the greater the similarity among the items to be learned, the greater was the number of errors, and the frequency of errors increased in spite of the repeated

readings which afforded additional opportunities for learning and corrections.

8. Poor students (as rated by respective class standing) underlined many more of the judgments than the better students, and the proportion of error among these alleged judgments based upon certainty, was always markedly larger among the poor students than among the pupils of high ranking in the classes.

A COMPARISON OF MARKS BASED UPON WEIGHTED AND UNWEIGHTED ITEMS IN A NEW-TYPE EXAMINATION

EDWARD F. POTTHOFF

University of Illinois

AND

NEVILLE E. BARNETT

Franklin, Ill. High School

The new-type examination is frequently designated by the term objective, and the belief is sometimes held that this type of examination, or at least certain forms of it, are objective in every respect. It is characteristic of the new-type test that the answers which a pupil gives to the items comprising it are either right or wrong, and not partly so. When, however, an attempt is made to determine the relative weights or values of the different items, the subjective element is introduced. The study here reported was carried out in the attempt to find out what effect the weighting of the items in a new-type test would have upon the marks, or literal grades, of individual pupils.

A test in American History, consisting of forty completion items and sixty multiple-choice items, was used for purposes of this investigation. The items comprising the test were weighted by ten high-school teachers of history, the instructions to the teacher directing him to select any item which in his opinion had a value of one as far as importance was concerned, and to weight the remaining items by comparing them with this standard. Consequently there were available ten sets of weights for the test, as well as the set in which the items were all weighted equally by assigning a value of one to each item. These sets of weights are referred to hereafter as scoring systems because each constituted a system by which the paper of a pupil taking the test could be scored.

The test was next administered to four hundred high-school pupils who were studying American History. The paper of each pupil was scored in each of the eleven ways referred to in the preceding paragraph. That is, there were obtained for each pupil eleven different scores, one based upon the scoring system supplied by each of the ten teachers and one derived by assigning equal values to all items—the unweighted method.

The scores which pupils earned under each of the eleven scoring systems were converted into literal grades by assuming that such grades would form a normal distribution, and that a marking system consisting of five categories, A, B, C, D, and E, would be used. It was further assumed that these marks would be distributed in the following manner: seven per cent A, twenty-four per cent B, thirty-eight per cent C, twenty-four per cent D, and seven per cent E. Consequently there were available for each pupil, eleven marks or letter grades, ten of the marks being weighted and one unweighted. These various methods of obtaining a pupil's mark are referred to hereafter as marking systems.

One of the problems undertaken in this study was that of determining the extent to which each of the teachers agreed with the unweighted marking system with respect to the marks assigned to the pupils' papers. That is, an attempt was made to find out how the papers which were placed in each letter category under the unweighted system were distributed among the various categories when marked by each of the teachers. Table I is introduced in answer to this question. The table shows, for example, that of all of the papers receiving an unweighted grade of B, 4.3 per cent received a grade of A, 82.8 per cent a grade of B, and 12.9 per cent a grade of C, when marked by teacher *Q*.

According to Table I the average agreement for all teachers and all categories of grades was eighty-eight per cent, the disagreement being divided equally between grades displaced one letter upward and those displaced one letter downward. On the average, therefore, the teachers disagreed with the unweighted system with respect to the marks assigned to twelve per cent, or forty-eight, of the four hundred papers concerned. This disagreement means that if a pupil's paper were graded by a weighted rather than by an unweighted system, he had, on the average, about one chance in eight of having his mark raised or lowered by one letter grade. In other words, the subjective element enters to a certain extent into the marking of papers on a new-type test when the matter of weighting the items comprising it is introduced.

The table shows further that some of the teachers agreed much more closely with the unweighted system than did others, the disagreement ranging from one per cent to twenty per cent when single categories are considered, and from 6.4 per cent to eighteen per cent when all categories are combined. The general agreement was lowest in the case of

TABLE I.—SHOWING BY PERCENTAGES HOW THE PUPILS PLACED IN EACH CATEGORY UNDER THE UNWEIGHTED MARKING SYSTEM WERE DISTRIBUTED AMONG VARIOUS CATEGORIES WHEN MARKED BY EACH TEACHER

Teacher	Unweighted mark																Agreements	One letter lower	One letter higher	Total
	A			B			C			D			E							
	Teacher's mark			Teacher's mark			Teacher's mark			Teacher's mark			Teacher's mark							
	A	B	A	B	C	B	C	D	C	D	E	D	E							
Q	86.7	13.3	4.3	82.8	12.9	8.0	84.0	8.0	12.4	81.4	6.2	20.0	80.0				83.0	8.5	8.5	100
R	86.7	13.3	4.3	89.2	6.5	4.0	89.3	6.7	10.3	85.6	4.1	13.3	86.7				88.0	6.0	6.0	100
S	86.7	13.3	4.3	81.3	14.0	8.7	82.0	9.3	14.4	80.2	5.4	16.7	83.3				82.0	9.0	9.0	100
T	93.3	6.7	2.2	90.3	7.5	4.7	90.0	5.3	8.3	89.6	2.1	6.7	93.3				90.4	4.8	4.8	100
U	93.3	6.7	2.2	91.3	6.5	4.0	90.7	5.3	8.3	88.6	3.1	10.0	90.0				90.4	4.8	4.8	100
V	93.3	6.7	2.2	91.3	6.5	4.0	90.0	6.0	9.3	87.6	3.1	10.0	90.0				90.0	5.0	5.0	100
W	83.3	16.7	5.4	88.1	6.5	4.0	88.0	8.0	12.4	83.5	4.1	13.3	86.7				86.4	6.8	6.8	100
X	80.0	20.0	6.5	91.7	11.8	7.3	86.0	6.7	10.3	84.5	5.2	16.7	83.3				84.0	8.0	8.0	100
Y	90.0	10.0	3.3	91.4	5.4	3.3	92.7	4.0	6.2	92.8	1.0	3.3	96.7				92.4	3.8	3.8	100
Z	93.3	6.7	2.2	93.5	4.3	2.7	94.0	3.3	5.2	92.7	2.1	6.7	93.3				93.6	3.2	3.2	100
Average.....	88.7	11.3	3.7	88.1	8.2	5.1	88.6	6.3	9.2	87.0	3.8	11.7	88.3				88.0	6.0	6.0	100

teacher *S* and highest in the case of teacher *Z*, amounting to 82.0 per cent and 93.6 per cent respectively for these two teachers. The coefficients of correlation between the unweighted scores and the scores assigned to the papers by these two teachers were .9655 ± .0022 and .9875 ± .0008 for teachers *S* and *Z* respectively. It appears, therefore, that weighted letter grades may disagree considerably with unweighted grades, even though the weighted scores and the unweighted scores are very highly correlated. In other words, although the weighting of the items in a test has only a slight effect upon the relative values of the scores which pupils make on such tests, nevertheless the change in the ranks of some of the scores may have considerable effect upon the marks or literal grades into which such scores are converted.

The correlation coefficients which have just been mentioned lead one to suspect that the differences between the marks assigned to papers under the unweighted and any one of the weighted marking systems were in reality very slight, even though this difference sometimes appeared as an entire letter grade. It may be, for instance, that the papers which received an unweighted mark of B and which were given a grade of C by teacher *Q* started out as very low B's under the unweighted system and ended up as very high C's under teacher *Q*.

In order to test this hypothesis the following study was made. The raw unweighted score assigned to each of the four hundred papers was converted into the corresponding standard score. The same thing was done with the raw scores assigned by teacher *Q* in order to secure two sets of scores which were directly comparable. The standard score given to each paper by teacher *Q* was then subtracted from the standard score assigned to it under the unweighted system. The difference in each case represented the extent to which the standard score changed from one scoring system to the other. These differences have been tabulated in Table II.

The data in this table may be interpreted more clearly by pointing out that one sigma distance represents almost exactly the distance across any letter grade and roughly the distance from the mean of one letter-grade category to the mean of the adjacent category. This relationship follows from the two assumptions which were made earlier with respect to the marking system: First that it consisted of five categories; and second, that it gave a distribution containing seven per cent each of A's and E's, twenty-four per cent each of B's and D's, and thirty-eight per cent of C's. When marks are distributed in this manner among the five categories, the distance on the base-line across

any category is approximately equal to one sigma, or one standard deviation.

This interpretation of the units of measurement employed in Table II makes it clear that in no case did a mark assigned to a paper by teacher *Q* actually differ from the unweighted mark by as much as a full letter-grade, the greatest difference being less than nine-tenths letter grades. Furthermore, a very large majority of the differences were small, being less than .5 sigma distance, or half of a letter-grade, in case of 93.5 per cent of the papers. Even though the differences between these weighted and unweighted scores were small, numerous papers received different marks, or letter grades, under the two systems of marking because even a very slight difference in the scores assigned to a paper by the two systems would make it possible for a mark which occupied a very high or a very low position in a given category under one system to be changed to a very low or a very high position in the adjacent category under the other system.

TABLE II.—SHOWING THE DISTRIBUTION OF THE DIFFERENCES OBTAINED BY SUBTRACTING THE STANDARD SCORE ASSIGNED TO EACH PAPER BY TEACHER *Q* FROM THE STANDARD SCORE ASSIGNED TO IT UNDER THE UNWEIGHTED SYSTEM

Differences in terms of standard scores or sigma distances	Number	Per cent	Cumulative per cent
.00 to .099	124	31.00	31.00
.10 to .199	108	27.00	58.00
.20 to .299	73	18.75	76.25
.30 to .399	47	11.75	88.00
.40 to .499	22	5.50	93.50
.50 to .599	18	4.50	98.00
.60 to .699	4	1.00	99.00
.70 to .799	3	.75	99.75
.80 to .899	1	.25	100
Total.....	400	100	

The extent to which various numbers of teachers agreed with the unweighted marking system with respect to the marks assigned to pupils' papers is shown in Table III. It indicates that all of the ten teachers agreed with the unweighted system upon the marks given to two hundred twenty-six of the four hundred papers. That is, there was unanimous agreement upon 56.50 per cent of the marks assigned.

In addition there were fifty-one papers which received the same marks from nine of the ten teachers as they did under the unweighted system. The last two columns of the table are cumulative, that is, they show the number and per cent of papers whose weighted marks agreed with the unweighted marks when the former were assigned by the number, or more than the number, of teachers indicated at each point. Thus, six or more teachers agreed with the unweighted marks given to three hundred sixty, or ninety per cent, of the papers. The table gives further indication of the fact that a considerable degree of subjectivity may be introduced into the marking of papers on a new-type test when attempts are made to weight the items comprising such a test.

Before concluding this discussion, attention should be called to the bearing of the reliability of tests upon the matter under consideration. It has already been stated that the correlations of weighted scores and unweighted scores were .9655 in the case of one teacher and .9875 in the case of another. These correlations are undoubtedly higher than the reliability coefficient of the test. That is, if the test had been administered a second time to the four hundred pupils shortly after they had taken it for the first time, the correlation between the pupils' scores on

TABLE III.—EXTENT TO WHICH VARIOUS NUMBERS OF TEACHERS AGREED WITH THE UNWEIGHTED MARKING SYSTEM WITH RESPECT TO THE MARKS ASSIGNED TO PAPERS

Number of teachers who agreed with the unweighted system	Papers agreed upon			
	Number	Per cent	Cumulative	
			Number	Per cent
Ten	226	56.50	226	56.50
Nine	51	12.75	277	69.25
Eight	42	10.50	319	79.75
Seven	21	5.25	340	85.00
Six	20	5.00	360	90.00
Five	19	4.75	379	94.75
Four	6	1.50	385	96.25
Three	5	1.25	390	97.50
Two	4	1.00	394	98.50
One	3	.75	397	99.25
None	3	.75	400	100
Total	400	100		

the first and second testings would be considerably less than .96. In other words, the variation introduced into pupils' scores on the test by weighting the items is smaller than the error which results from the fact that the test is not entirely accurate as a measuring instrument. Consequently it follows that for practical purposes the differences between weighted and unweighted scores may be considered to be so small that they may be disregarded, and a great deal of labor may be dispensed with by using the unweighted scores in determining the literal grades. Nevertheless the fact remains that weighted and unweighted scores lead to different literal grades on some papers, and that the new-type examination is not found to be completely objective when the question as to the values or weights of the items composing it is considered.

ITEMS OF GESELL'S DEVELOPMENTAL SCHEDULE SCALED

MINNIE L. STECKEL

University of Chicago

The object of this study was to apply the method of absolute scaling, as devised by Thurstone[1] to a set of data covering very young age groups in an effort to determine the type of distribution of ability, the absolute zero of intelligence, and the shape of the mental growth curve when the ratings approach very near the time of birth; also to provide a comparative rating scale for children of very young age groups.

The study is based upon the items of Gesell's[2] Developmental Schedule. This schedule was selected because it includes ratings on very low age groups which were obtained from a representative sample of pre-school children.

In applying the absolute method of scaling to a set of data only those items can be utilized for which there are ratings on the same basis in each of two or more adjacent age groups. In applying this method of scaling to Gesell's Developmental Schedule it was necessary, therefore, to eliminate all items which occurred as a rating in only one age group. The first task was to select and tabulate all items for each age period which occurred in any two adjacent age groups. There were from eleven to twenty-four overlapping items for each of the various age groups. While it would be more desirable to have a greater number of overlapping items to base the study on, the results indicate that these are sufficient to make the conclusions valid.

The Gesell ratings, $a+$, a, $b+$, b, c, are applied to percentage groups; one per cent to nineteen per cent, twenty per cent to forty-nine per cent, fifty per cent, to sixty-four per cent, sixty-five per cent to eighty-four per cent, eighty-five per cent to one-hundred per cent, respectively. In order to apply the method of absolute scaling to these data the mid-point of each percentage group was taken as the rating upon which to base the calculation.

---

[1] Thurstone, L. L.: A Method of Scaling Psychological and Educational Tests. *Journal Educational Psychology*, 1925, Vol. XVI, pp. 433–451.

[2] Gesell, Arnold: "The Mental Growth of Pre-school Child." The Macmillan Company, 1925.

Table I gives the data from the Gesell Schedule for the six-months and nine-months old children. In the first column is indicated the number of the item which was selected because it occurs in both of two adjacent age groups. In three cases the same number occurs twice. In the Gesell Schedule several items are listed under one serial number. In this study each item of the series is considered as a separate item and so is listed under the same number with a subscript to indicate the particular item of the series. The second column gives the description of the Gesell item. Columns three and four give the Gesell Ratings for the six-months old and nine-months old children respectively. In the next two columns are indicated the relative difficulty of each question for each age group expressed in terms of standard deviation. These figures are obtained by using the mid-point of the percentage ratings as recorded by Gesell for each age group. In the last two columns the squares of the sigma values from which the standard

TABLE I

Gesell item	Description of Gesell items	Gesell rating for six months	Gesell rating for nine months	Mid-point percentage of six-month-olds who answered correctly, P_6	Mid-point percentage of nine-month-olds who answered correctly, P_9	Difficulty for six-month-olds, X_6	Difficulty for nine-month-olds, X_9	X^2_6	X^2_9
*M*20	Creeps or hitches	*a*+	*a*	10	35	1.28	.39	1.64	.15
*M*25	Makes stepping movements	*a*	*b*	35	75	.39	− .67	.15	.45
*P*41	Bangs spoon	*b*	*c*	75	92	− .67	−1.41	.45	1.99
*M*34	Manipulates with one hand	*a*	*c*	35	92	.39	−1.41	.15	1.99
*A*31	Lifts inverted cup	*b*	*c*	75	92	− .67	−1.41	.45	1.99
*A*31$_2$	Secures cube	*a*+	*a*	10	35	1.28	.39	1.64	.15
*A*30	Looks for fallen object	*a*	*c*	35	92	.39	−1.41	.15	1.99
*M*35	Holds two cubes	*b*	*c*	75	92	− .67	−1.41	.45	1.99
*M*35$_2$	Drops for third	*b*+	*b*	57	75	− .18	− .67	.03	.45
*M*36	Releases cube in cup	*a*+	*a*	10	35	1.28	.39	1.64	.15
*A*15	Grasps above head	*b*+	*c*	57	92	− .18	−1.41	.03	1.99
*A*15$_2$	Persistent reaching	*a*	*b*	35	75	.39	− .67	.15	.45
*A*18	Purposeful reaction to paper	*a*	*b*	35	75	.39	− .67	.15	.45
*P*13	Reacts to mirror image	*a*	*b*+	35	57	.39	− .18	.15	.03
*P*40	Frolics when played with	*b*+	*c*	57	92	− .18	−1.41	.03	1.99
*L*10	Says mama, dada, or equivalent	*a*+	*b*	10	75	1.28	− .67	1.64	.45
*P*42	Reacts to music, coos	*a*	*b*+	35	57	.39	− .18	.15	.03
*M*32	Fine prehension	*a*+	*b*	10	75	1.28	− .67	1.64	1.45
*A*17	Exploratory manipulation	*a*	*b*	35	75	.39	− .67	.15	.45
*A*16	Reaches for spoon	*b*	*c*	75	92	− .67	−1.41	.45	1.99

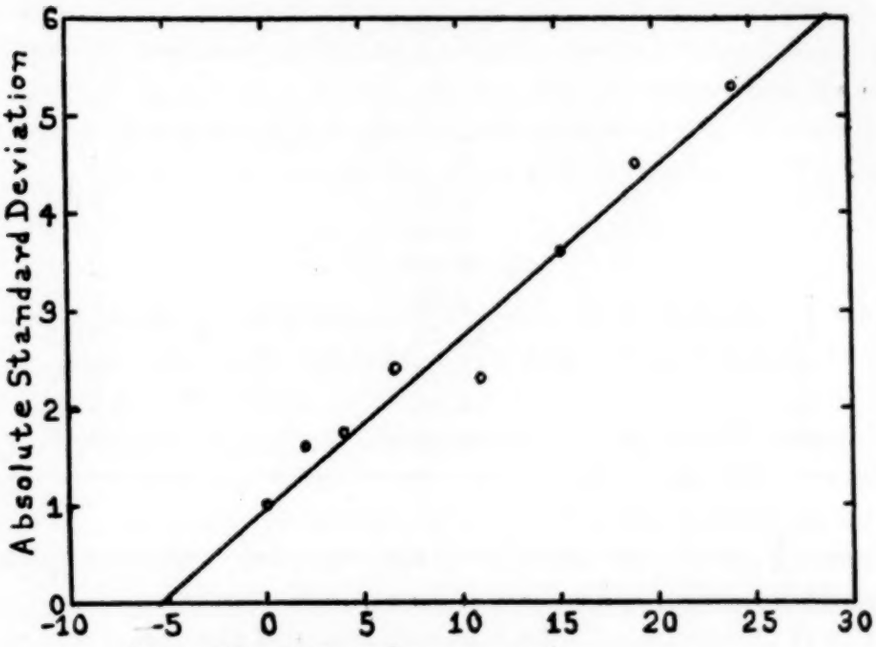


Figure I

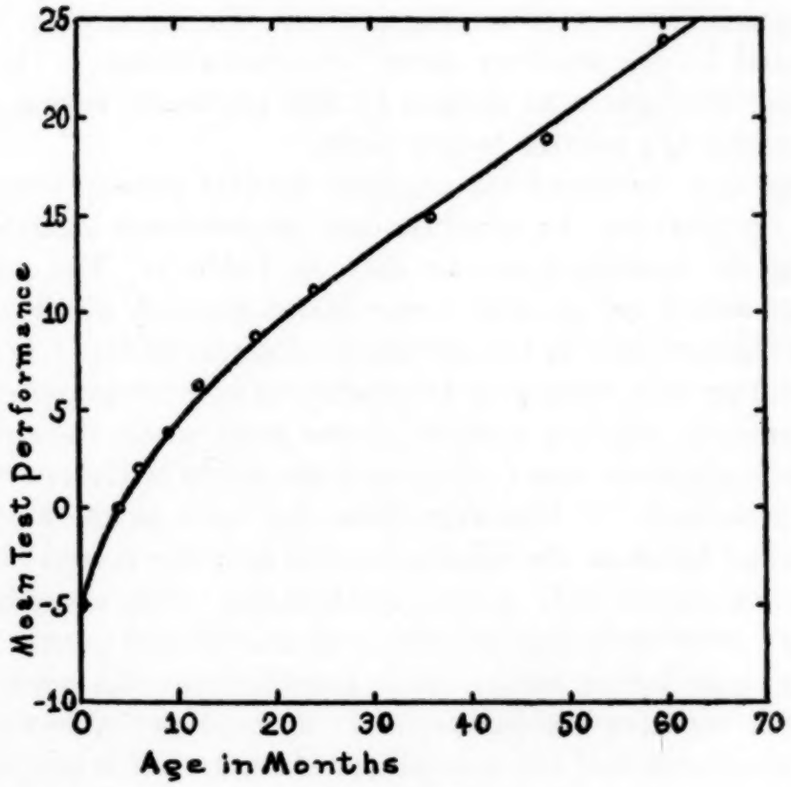


Figure II

TABLE II

Mean age in months	M	σ
4	0	1
6	2.01	1.63
9	3.84	1.75
12	6.41	2.27
18	8.89	1.73
24	11.15	2.32
36	14.74	3.62
48	18.84	4.50
60	23.99	5.28

deviations S_6 and S_9 are calculated are recorded. Similar calculations were made for every other two adjacent age groups.

Table II shows the absolute variability and the mean test performance of each group for the age range four months to sixty months inclusive.

Figure 1 shows the linear plot of σ against M. The straight line is fitted to the data by the method of least squares and verifies the law of variability of test intelligence *i.e.*, the variability is directly proportional to the absolute mean test performance. The absolute zero of test intelligence as defined by this particular rating scale is at approximately $5\frac{1}{2}$ months before birth.

In Fig. 2 is indicated the absolute mental growth curve as constructed by plotting the absolute test performance against chronological age in months from the data in Table I. The point $-5\frac{1}{2}$ months at which the growth curve passes through absolute zero, as previously determined by the variability diagram of Fig. 1, is indicated. The mental growth curve may be completed by extrapolating from the lowest age norm, age four months, to the point where the curve passes through the absolute zero ($-5\frac{1}{2}$) and the shape of the curve near the origin determined. If this were done the error in the curve should be very small for since the data extend so near the period of birth the extrapolation covers only a very small range. The growth curve as determined from these data is negatively accelerated as near the origin as four months before birth. It is possible that the curve actually begins with negative acceleration. It is more probable that at the union of the sperm and ovum intelligence begins but is not perceptible until four or five months before birth. It seems reasonable to con-

jecture that at the union of sperm and ovum there is a positive acceleration which the trauma at birth changes to negative acceleration which continues through-out the rest of the life period. For this reason the shape of the curve from the absolute zero to the lowest point for which we have data is still in question and is left to the reader's own discretion.

Conclusions

A linear relation is found to exist between the absolute variability and the mean test performance of successive age groups of the test items scaled in this study.

The distributions of test ability of the adjacent age groups are normal.

The absolute zero of intelligence for this particular set of data is defined at $-5\frac{1}{2}\sigma$.

The mental growth curve for the age range covered is negatively accelerated.

EYE MOVEMENTS AND THE EYE-HAND SPAN IN TYPEWRITING

RUSSELL L. C. BUTSCH

Marquette University

Purpose of the Study

The purpose of this experiment is to determine the way in which the eye follows the copy while one is typewriting, and especially to discover the distance which the eye keeps ahead of the hands in writing at various speeds. This investigation is analogous to that carried on by Buswell[1] in which he studied the eye-voice span in oral reading. The span in the case of typewriting might be called the "eye-hand span," and that term will be used in this report.

The investigation by Buswell mentioned above was carried on by means of the apparatus for photographing eye movements, and a synchronized dictaphone. In that study it was found that the slow and inefficient readers read practically a syllable at a time, and their eyes did not get very far ahead of the voice. But the efficient and rapid readers were found to have an eye-voice span as wide as forty or more letter spaces in some cases, and the average span for good readers among adults was close to twenty letter spaces.

There have been a few other experiments which attempted to determine either directly or indirectly the amount of time elapsing between the stimulus and the response in the case of complex activities. In general the conclusions have been similar to those reported by Buswell for oral reading, namely, that as efficiency increased there was a continual increase in the length of the span until at the level of highest efficiency the response was separated by a considerable interval from the particular stimulus calling it out. Bryan and Harter, in their studies on the telegraphic language, found that as the subjects improved in their receiving speed there was a constantly widening span between the hearing of the letter and its transcription. They state:

> The beginner must take each letter as it comes, *i.e.*, he can copy behind one letter. Later he can wait for words. A fair operator can copy behind several

[1] Buswell, Guy Thomas: An Experimental Study of the Eye-voice Span in Reading. *Supplementary Educational Monographs, No.* 17. Chicago; Department of Education, University of Chicago, 1920.

words in connected discourse. The expert prefers to keep six to ten or twelve words behind the instrument.[1]

Book, in his study, found exactly the same results for typewriting that Bryan and Harter had found for the telegraphic language. He says:

> The copy in the word association stage is taken as a syllable or as a word at a time and a little ahead of the hands.[2]
>
> In expert stage word and phrase associations have been developed to such a degree that the writing is absolutely continuous. The word phrase and clause have become the unit of attention . . . Attention was pushed ahead of the hands as far as possible (usually four or five words).[3]
>
> The copy in this stage is gotten by reading it a word at a time, but, as has been said, a number of words ahead of the hands.[4]

These statements by Book are based more on the introspections of his subjects than on any objective evidence. The present study attempts to determine objectively the length of the eye-hand span in typewriting at different speed levels, and with different types of copy. The technique is similar to that used by Buswell in his study of oral reading.

The Apparatus

The apparatus used was the camera for photographing eye movements, in the educational laboratory of the University of Chicago, with an attachment for synchronizing the action of the typewriter with the film record. A few other slight modifications of the usual technique for photographing eye movements had to be adopted. Since the apparatus used for photographing eye movements has been fully described elsewhere,[5] the present description will cover only the modifications and additions necessary for this investigation.

One variation from the usual procedure was in the method of holding the subject's head to eliminate head movements. In the preliminary trials it was found that there was a great deal of such movement while the subject was typewriting. The ordinary head clamps

---

[1] Bryan, W. L. and N. Harter: Studies on the Telegraphic Language. *Psychological Review*, Vol. VI, p. 353, 1899.

[2] Book, W. F.: "The Psychology of Skill." P. 37.

[3] *Ibid.*, p. 44.

[4] *Ibid.*, p. 45.

[5] Gray, Clarence Truman: Types of Reading Ability as Exhibited through Tests and Laboratory Experiments. *Supplementary Educational Monographs*, Vol. I, No. 5, pp. 83–91. Chicago: University of Chicago, 1917.

are applied to the skin on the cheek bones and the skull, and thus permit movement of the skin over the bones. To eliminate this movement the following method was used. An impression of the subject's upper front teeth was obtained in a piece of sealing wax placed on the end of a small wooden paddle. This paddle was then clamped securely in place, and held tightly against the teeth. Since the upper jaw furnishes a point at which contact can be had directly with the bone of the skull this device, supplementing the usual head clamps, was found to be reasonably satisfactory.

A typewriter, fastened securely to a small, heavy table, was placed directly in front of the subject, at as nearly the correct height for typewriting as possible. Fastened by braces to the two ends of the carriage slide rod, and running parallel with the carriage at the distance of about two inches behind the typewriter, was a contact rod. The upper surface of this rod was composed mainly of bakelite, except that there were narrow strips of brass running across the rod, one inch apart. The lower part of the rod was all of brass, thus making a connection through any one of the narrow strips to the binding post fastened to the end of the rod. Fastened to the right side of the typewriter carriage there was a small brass rod about three inches in length, which pressed down vertically against the contact rod. This latter short rod moved with the carriage, while the longer contact rod remained stationary. There was also a binding post on the end of the short vertical rod. When the two binding posts were connected with a dry cell and the typewriter carriage was moved from one side to the other, the circuit was closed each time the vertical rod touched one of the narrow strips of brass on the contact rod. These strips were so placed that the circuit was closed every time the carriage moved forward ten spaces. An ordinary electro-magnetic time marker was connected in circuit with the contact apparatus and several dry cells, and set in such a way that each time the circuit was closed the marker cut off the light which, after reflection from the subject's eye, produced the record on the film.

Materials and Procedure

The selections used in this experiment were typewritten on white cardboard. The procedure was the same as in the experiments in which the eye-movements in reading are photographed, except that instead of merely reading silently or orally, the subject reproduced on the typewriter the material in the copy.

Subjects

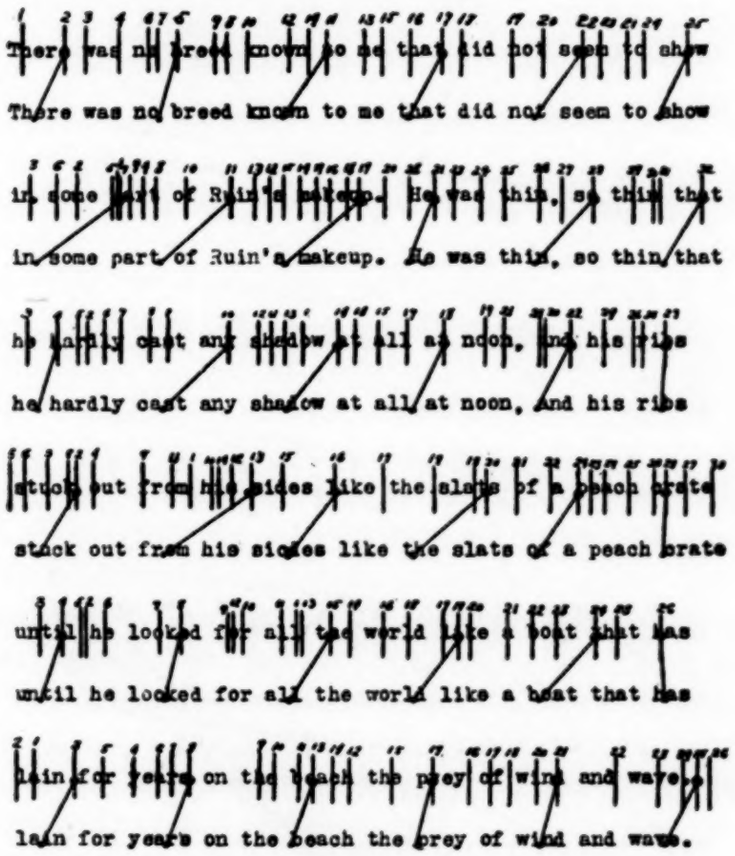
The subjects used in this experiment were obtained through the courtesy of a number of individuals and institutions. Two Chicago business colleges, the Bryant and Stratton Business College and the Gregg School very kindly cooperated by sending some of their students at various stages of progress as subjects. Other subjects were obtained from a class in the Teaching of Shorthand and Typewriting at the

There was no breed known to me that did not seem to show
in some part of Ruin's makeup. He was thin, so thin that
he hardly cast any shadow at all at noon. And his ribs
stuck out from his sides like the slats of a peach crate
until he looked for all the world like a boat that has
lain for years on the beach the prey of wind and wave.

Plate I.—Subject G. B. (Slowest Writer).

University of Chicago. Through the courtesy of the Underwood Typewriter Company, Mr. William F. Oswald, former world's champion typist, also served as a subject.

Nature of the Records

For each subject there was obtained a film which recorded the movement of the eye in reading the copy. Breaks in the line on the record indicated the location of the eye each time the contact points were touched. The machine was set in each case so that the first

break came after the writing of the second letter of the line, and before the third letter was struck, and the other breaks came at intervals of ten spaces thereafter. There were six breaks in each line.

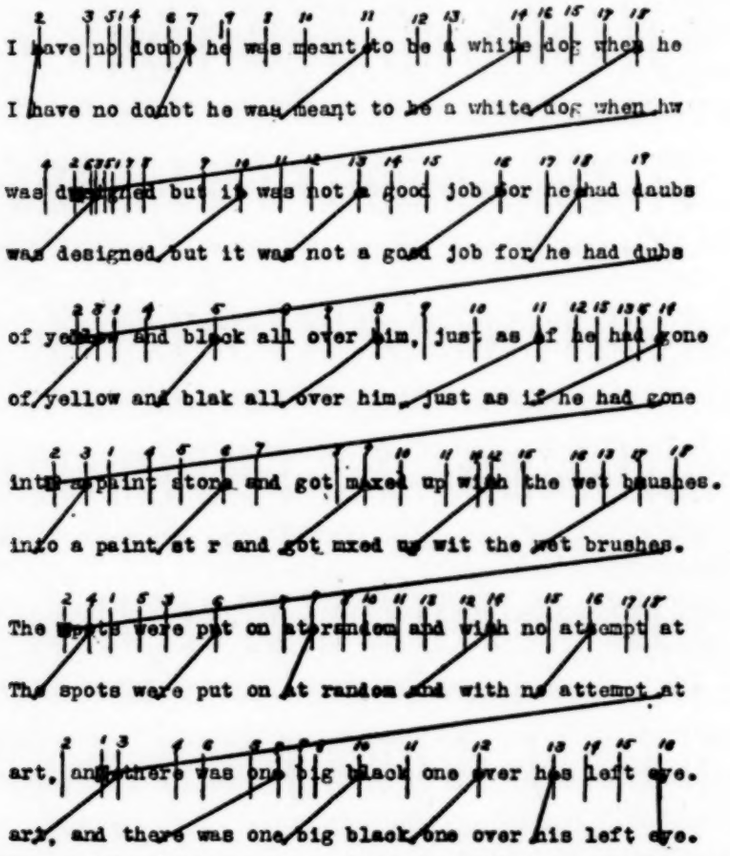
After the films were developed they were plotted in the usual manner. As the film was moved through the projection lantern each fixation was located on the proper line, and the order of these fixations for any particular line was indicated by the numbers placed above the fixations. In addition, the location of each break, or point where the

I have no doubt he was meant to be a white dog when he
I have no doubt he was meant to be a white dog when hw
was designed but it was not a good job for he had daubs
was designed but it was not a good job for he had dubs
of yellow and black all over him, just as if he had gone
of yellow and blak all over him, just as if he had gone
into a paint stone and got mixed up with the wet brushes.
into a paint st r and got mxed up wit the wet brushes.
The spots were put on at random and with no attempt at
The spots were put on at random and with no attempt at
art, and there was one big black one over his left eye.
art, and there was one big black one over his left eye.

PLATE II.—Subject E. F. (Most Rapid Student Writer.)

light was cut off, was also marked. This plotting was then transferred to a copy of the selection written. Plate I presents the plotting for the slowest writer included in the study, and Plate II for the most rapid student writer. In the case of each pair of lines, the first represents the copy from which the subject wrote, and the second represents what the subject wrote. In addition to the vertical lines indicating the fixations, small black discs are placed on certain of the fixations in the first of each pair of lines to indicate the points at which the subject was looking when the circuit was closed and the light shut off.

The dots in the second of each pair of lines indicate the position of the carriage as these breaks occurred. These were easily located since, as explained above, the number of spaces from the beginning of the line was known for each break. The diagonal lines connecting the dots and the discs are put in to show more clearly the distance between the eye and the hand at each break. A long diagonal line from the end of one line to the beginning of the next indicates that while the subject was writing a word at the end of the one line, his eyes had already moved on to the beginning of the next line.

Data for the Individual Subjects

From the plottings for each subject the following data of a statistical nature were obtained:

1. The number of fixations in each line; and the average number of fixations per line.
2. The number of regressive movements in each line; and the average number of regressive movements per line.
3. The number of spaces the eye was ahead of the hand on each break; and the average number of spaces the eye was ahead of the hand throughout the selection—the eye-hand span.
4. The number of times the eye moved on to the next line before the hand had finished writing the preceding line; and the percentage of the possible number of times that this occurred.

The data obtained for nineteen subjects are summarized in Table I. In the second column is indicated the speed at which the subject ordinarily wrote, in words per minute. In the succeeding columns are shown the average number of fixations per line, regressive movements per line, eye-hand span, and the percentage of cases in which the span was carried over the end of the line.

One of the first facts which is indicated in an examination of Table I is that while there is considerable difference between the slowest and the most rapid writers, so far as the eye movements are concerned, this difference is not consistent. One subject with a speed of fifty words per minute made an average of 17.8 fixations per line, while one with a speed of seventy-five words made an average of 17.5 fixations. Likewise the other items in the table exhibit a lack of agreement between subjects writing at the same rate, and a lack of consistent advance from the slower to the more rapid writers. However, in considering the table as a whole and following down through any column it is possible to observe a general trend. In general, the slower

writers require more fixations per line; in general, the eye-hand span is larger for the more rapid writers; in general, the more rapid writers carry the span over the end of the line in a larger percentage of the cases.

TABLE I.—STATISTICAL SUMMARY BY INDIVIDUAL SUBJECTS

Subject	Speed	Average number of fixations	Average number of regressive movements	Average "eye-hand span"	Percentage carried over end of line
G. B.	35	27.0	6.7	3.24	0
F. P.	40	26.2	3.8	4.64	40
E. M.	50	18.2	1.2	3.44	0
H. H.	50	17.8	1.4	3.33	50
R. C.	50	20.6	2.0	4.20	75
G. K.	55	19.0	2.8	3.90	75
R. D.	55	18.5	3.0	3.79	67
S. C.	55	24.0	6.0	4.50	75
E. D.	60	16.4	1.2	3.57	40
M. S.	60	18.0	2.3	4.46	67
O. D.	60	20.4	2.8	5.63	100
M. M.	60	16.7	2.3	5.64	80
M. E.	65	20.2	2.8	5.17	80
A. K.	65	15.7	2.2	5.55	60
C. N.	65	17.2	3.0	6.31	100
N. R.	70	15.3	0.7	4.69	40
G. M.	70	16.5	1.3	7.60	100
E. F.	75	17.5	3.0	5.53	100
W. O.	100	16.3	1.7	6.44	100

DATA BY SPEED GROUPS

The general trend of improvement from slower to more rapid writers is made more evident by an examination of Table II, in which the data are summarized by giving the averages of the various speed groups. The subjects writing forty words per minute or less are grouped together, and indicated by 40; those writing seventy and seventy-five words per minute are grouped together, and indicated by 70. The data for the one hundred word group in this table include two similar tests written by Mr. Oswald, instead of the single test included in Table I. In all terms of this table except that of regressive movements there is seen to be rather steady improvement from the lower to the higher speed groups.

One very striking fact brought out by both Table I and Table II is the large number of fixations per line required even among the more rapid writers. The smallest number of fixations recorded is an average of 15.3 per line, which means, in a line of sixty spaces, the inclusion of only 3.6 letters per fixation. When the number of fixations per line

TABLE II.—STATISTICAL SUMMARY BY SPEED GROUPS

Speed group	Number of cases	Average number of fixations	Average number of regressive movements	Average "eye-hand span"	Percentage carried over end of line
40	2	26.3	5.0	3.96	20
50	3	18.9	1.5	3.70	42
55	3	20.5	3.9	4.46	72
60	4	17.9	2.2	4.88	72
65	3	17.7	2.7	5.69	80
70	3	16.4	1.7	5.93	80
100	1	16.3	1.7	6.83	94

here reported is compared with the results given by Buswell for lines of approximately equal length—5.9 fixations for silent reading, and 8.4 for oral reading[1]—it is seen that the eye is not simply reading when it is obtaining copy for typewriting. A comparison of the films obtained during typewriting, and films obtained for reading for the same subjects, reveals the fact that not only are the fixations more numerous but they are also of longer duration. Apparently during typewriting the eye does not read at all at its maximum pace or even at the rate which is determined by the requirements of comprehension, but instead reads only rapidly enough to supply the copy to the hand as it is needed.

THE LENGTH OF THE EYE-HAND SPAN

The data on the eye-hand span in Tables I and II indicate that the average length for individual writers varies from 3.24 spaces to 7.60 spaces, and the averages for groups vary from 3.70 spaces to 6.83 spaces. These averages are to be contrasted with the eye-voice span reported by Buswell of twenty or more spaces on the average, with some running as high as forty spaces. They should also be contrasted

[1] Buswell, G. T.: *Op. cit.*, pp. 26, 34.

with the statement by Book, quoted above, that expert typists read four or five words ahead of their writing.

The variation in the span in individual records and in records within the speed groups is equally as important as the averages. Table III presents the distributions of the spans for all points at which they were determined, by speed groups. There are in general six such points to the line, or thirty-six for each test. However, in some cases parts of tests could not be used, resulting in a number of cases not a multiple of six. An examination of this table indicates that the averages are not at all significant of the limits; the eye may often be less than three or more than seven spaces ahead of the hand. A total of twenty-two cases are found in which the eye was more than ten spaces ahead. In twenty-three cases the eye was on the exact letter that was being struck, and in twenty-seven cases only one letter ahead. There is, however, very clear evidence in this table that there is a tendency for the width of the eye-hand span to vary directly with the speed of the

TABLE III.—DISTRIBUTIONS OF LENGTHS OF EYE-HAND SPAN BY SPEED GROUPS

Length of span	Number of cases in each speed group						
	40	50	55	60	65	70	100
13	.....	.....	.....	1			
12	.....	.....	.....	3	2	3	2
11	.....	1	.....	2	4	1	3
10	.....	.....	1	2	5	6	7
9	1	.....	1	3	4	10	9
8	4	1	5	12	8	7	10
7	4	3	7	4	14	16	15
6	7	16	11	17	19	16	10
5	11	18	20	18	14	19	16
4	12	19	17	12	19	9	6
3	13	23	13	27	10	9	5
2	15	10	8	9	2	6	1
1	2	9	4	5	5	3	
0	3	8	3	5	2	2	
Total.....	72	108	9	120	108	107	84
Mean.....	3.96	3.70	4.46	4.88	5.69	5.93	6.83
PE_M.....	.16	.13	.15	.17	.17	.17	.17
SD.....	2.05	2.01	2.05	2.78	2.60	2.63	2.28

writing. There is also evidence that writers at any of the speeds included in the study do on occasion make use of spans much larger than the mean, and presumably could do so regularly if there were any need for it. It is probable that the eye simply keeps far enough ahead to provide copy for the hand as it is needed, and rapidly enough to prevent retarding the writing.

Significance of the Difference between the Various Speed Groups in Eye-hand Span

In Table III there are included the means and the probable errors of the means of the eye-hand spans for all speed groups. The question arises as to whether the differences between these means indicate real progress as speed increases, or whether they are the result of chance errors of sampling. This question may be answered by determining the statistical significance of the differences between these means. The difference and the probable error of the difference was computed for each pair of speed groups. It was found that there is a statistically significant difference between the spans for any given speed group and all groups separated from it by a difference of ten words a minute, except in the case of the three lowest groups. The data for the two slowest groups (fifty and forty) are so contradictory that one is led to assume either that the speeds reported were incorrect, or that the subjects were so inexpert that they were confused by the conditions of the experiment. In general the conclusion may be stated that an increase in speed of ten words a minute will be accompanied by a significant increase in the length of the eye-hand span.

The Time Interval Represented by the Eye-hand Span

Unfortunately the time required to write the tests used in this study was not recorded. However, the writing speed reported for the subjects may be used as a rough approximation. It is interesting to attempt to discover the time interval represented by the eye-hand span at different speeds. The speeds set down in the tables are in five-stroke words per minute. It is, therefore, easy to compute the number of strokes per second, which would be simply one-twelfth of the number of words per minute. These are computed for each speed group, and are set down in Table IV in comparison with the average eye-hand span for each group. An examination of this table indicates

TABLE IV.—RELATIONSHIP BETWEEN STROKES PER SECOND AND AVERAGE EYE-HAND SPAN FOR THE VARIOUS SPEED GROUPS

Speed in five-stroke words per minute	Speed in strokes per second	Average eye-hand span in strokes	Average number of seconds in each span
40	3.33	3.96 ± .16	1.19
50	4.16	3.70 ± .13	.89
55	4.58	4.46 ± .15	.97
60	5.00	4.88 ± .17	.98
65	5.42	5.69 ± .17	1.05
70	5.83	5.93 ± .17	1.02
100	8.30	6.83 ± .17	.82

a relationship between these two factors which is too nearly uniform to be explained solely as the result of chance. In five groups the ratio between the strokes per second and the eye-hand span is approximately 1.00. In four groups out of the seven the two are equivalent within the probable error. It is likely that the discrepancies of the data are to be explained as in the last section. The evidence, based of course on only a few cases, seems to be that the eye keeps at an average distance ahead of the hand such that the time interval between seeing a letter and writing it is approximately one second, no matter what the speed of the writing.

The conclusion suggested in the above paragraph applies, of course, only to the average condition. It is also clear from Table III that in exceptional cases the time interval may be much greater. For example, in the group writing at sixty words per minute, or five strokes per second, one eye-hand span of thirteen spaces is found. This would indicate an interval of 2.6 seconds. In the group writing forty words per minute, the longest span indicated would be 2.7 seconds. In the group writing at a rate of seventy words, the longest time interval would be 2.1 seconds. On the other hand, the cases at the other end of the distribution indicate that at times there was a very negligible time interval. It should be pointed out that, among the more rapid writers, these short spans occurred mainly at the beginning of the first line and at the end of the last. Among the slower writers they might be found at almost any point in the writing. There is no way of knowing whether or not the writing was proceeding at the normal speed at these points.

Writing by Phrases

In order to determine to what extent the writing was carried on by phrases at the different speed levels, a tabulation was made of the distance, in words, between the eye and the hand. These data are presented in Table V. The averages for the forty and fifty word groups are .83 and .76, respectively. This indicates that in approximately three-fourths of the cases the eye was resting on the following word. For no group does the average reach 1.50, except for the former champion. For all other groups the indication is that, on the average, the eye is on the second word following the one being written in less than half of the cases. In other words, although there is an increase in the average span, in terms of words, as speed increases, even the

Table V.—Distribution of the Number of Words Separating the Eye-fixation and the Hand, by Speed Groups

Span in words	Speed group						
	40	50	55	60	65	70	100
4	.....	.....	.....	1			
3	1	.....	.....	8	7	6	9
2	9	10	16	23	33	39	27
1	37	57	50	53	48	35	32
0	23	34	19	28	8	14	5
Total.....	70	101	85	113	96	94	73
Mean.....	.83	.76	.97	1.12	1.41	1.39	1.55

most rapid writers here included are not in the habit of reading two words in advance of the hands more than half of the time. The ordinary situation, with all except the most rapid writers, seems to be to write a particular word as the next word is read.

The cases in which the eye was found to be more than two words ahead of the hand are important as giving evidence of the writing by phrases. The longest such phrase recorded was found in the case of a subject in the sixty word group. It is: "him, just as if he." As the subject was writing the word "him"—to be exact, just as he struck the "m"—his eye was resting on the first letter of the word "he." In this case the eye was thirteen spaces ahead of the hand. More common examples of longer phrases are as follows:

"for he had daubs" (four subjects—span of 12 in 2 cases)
"if he had gone" (two subjects—span of 12)
"were put on at" (two subjects—span of 11)

Other four word phrases that were found are as follows: "but it was not"; "him, just as if"; "good job for he"; "got mixed up with"; "of a peach crate"; "was not a good"; "there was one big"; "into a paint store"; "known to me that"; "as if he had." The last one in the list is the only case of a phrase in which the eye was three words ahead of the hand in the records of the students writing forty words per minute or less.

The three word phrases, in which the eye was two words ahead of the hands, are much more numerous and varied, but of about the same general type. In most cases they are phrases conveying a unit of thought. The most popular one was "be a white," which was found in the records of three subjects in the seventy word group, two in the sixty-five, two in the sixty, and one in the forty word group. Others that were found on more than one paper were: "for he had"; "mixed up with"; "boat that has"; "was not a"; "the wet brushes"; "and his ribs"; "there was one"; "if he had"; "one big black." In some cases the three words were not a unit, indicating that the span may occasionally be due not to the sense, but to the eye habits, or to accidental circumstances. Samples of this sort are: "over him, just"; "white dog when"; "art, and there"; "store and got"; "random and with"; "good job for." However, in general, most of the phrases were made up of words which conveyed a unit of thought, and of small, commonly used words.

The Effect of Variation in the Type of Material

Records were obtained for Mr. Oswald, the former world's champion typist, on five selections. Tests I and V were of easy materials; Test II was made up of alphabetic sentences—that is, sentences each of which contain every letter of the alphabet; Test III was made up of a series of disconnected words of different lengths from three to seven letters; and Test IV was made up of rather difficult material. The results on these five tests are presented in Table VI. The data are also given separately for the words of different lengths in the selection made up of disconnected words. An examination of Table VI indicates that the number of fixations per line is practically the same for all material, varying from 15.8 to 17.1. Strangely enough it is least on the disconnected words, and largest on the easy selection

TABLE VI.—STATISTICAL DATA FOR MR. OSWALD FOR DIFFERENT TYPES OF MATERIAL

Type of selection	Average number of fixations	Average number of regressive movements	Average eye-hand span	Percentage carried over end of line
I. Easy material	16.3	1.7	6.50	100
II. Alphabetic sentences	16.6	1.7	7.71	100
III. Disconnected words	15.8	1.1	6.59	100
Three letter	17.0	0.5	7.33	100
Four letter	15.5	2.0	7.25	100
Five letter	15.7	1.0	6.11	100
Six letter	14.0	0.5	7.17	100
Seven letter	17.0	1.5	5.67	100
IV. Difficult material	17.1	1.8	6.17	67
V. Easy material	17.0	1.8	7.08	86

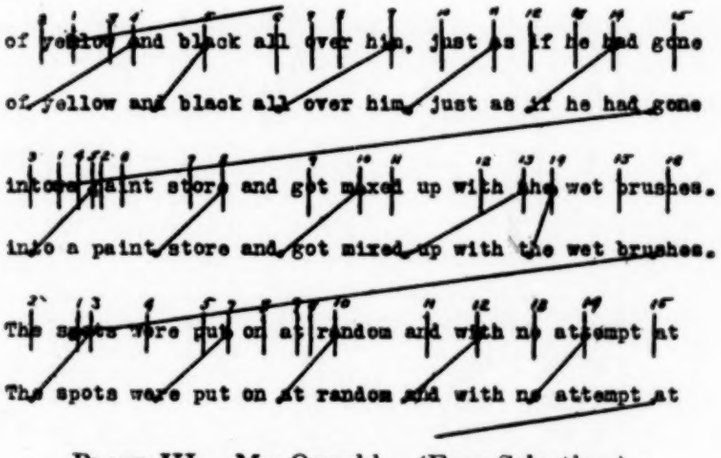

of yellow and black all over him, just as if he had gone
of yellow and black all over him, just as if he had gone
into a paint store and got mixed up with the wet brushes.
into a paint store and got mixed up with the wet brushes.
The spots were put on at random and with no attempt at
The spots were put on at random and with no attempt at

PLATE III.—Mr. Oswald. (Easy Selection.)

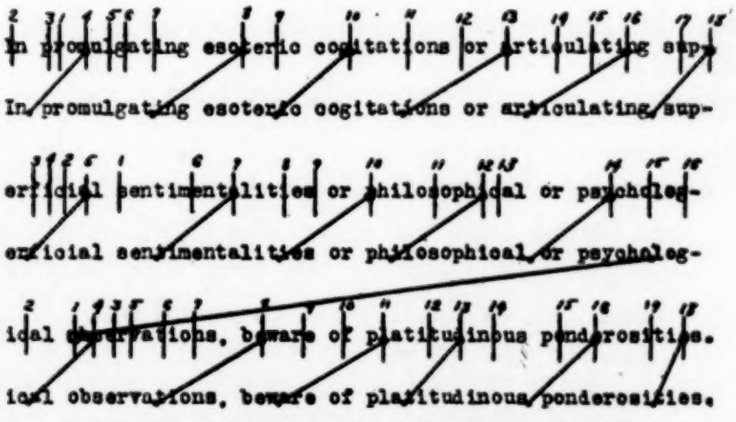

In promulgating esoteric cogitations or articulating sup-
In promulgating esoteric cogitations or articulating sup-
erficial sentimentalities or philosophical or psychaleg-
erficial sentimentalities or philosophical or psychaleg-
ical observations, beware of platitudinous ponderosities.
ical observations, beware of platitudinous ponderosities.

PLATE IV.—Mr. Oswald. (Difficult Selection.)

written last. The eye-hand span varies from 6.17 to 7.71. The plottings for parts of one easy selection and of the difficult selection are presented in Plates III and IV.

The same analysis has been applied to the scores of Mr. Oswald on the various types of tests as was used to contrast the records for the various speed groups. The distribution of the length of the eye-hand span in strokes for the different tests are presented in Table VII. The

TABLE VII.—DISTRIBUTION OF LENGTHS OF EYE-HAND SPAN ON THE VARIOUS TESTS TAKEN BY THE FORMER WORLD'S CHAMPION

Length of span	Test number									
	I	V	II	IV	III	(3)	(4)	(5)	(6)	(7)
12	.....	2	.....	.....	1	.....	1			
11	.....	3	3	.....	1	.....	1			
10	2	5	4	1	3	1	.....	1	1	
9	4	5	6	4	6	3	.....	.....	2	1
8	6	4	10	4	9	2	3	2	2	
7	6	9	9	6	14	2	2	5	2	3
6	5	5	5	7	13	2	2	5	2	2
5	8	8	4	7	9	1	2	2	.....	4
4	3	3	1	5	7	1	1	1	3	1
3	2	3	.....	1	1	.....	.....	1		
2	.....	1	.....	1	2	.....	.....	1	.....	1
Total	36	48	42	36	66	12	12	18	12	12
Mean	6.50	7.08	7.71	6.17	6.59	7.33	7.25	6.11	7.17	5.67
PE_M	.21	.24	.18	.21	.17	.34	.45	.29	.39	.33
SD	1.88	2.51	1.73	1.84	2.01	1.75	2.28	1.82	1.99	1.70

distributions for the words of varying length in Test III are also given separately.

The differences between the average lengths of the eye-hand spans for the various tests are not particularly striking. The significance of these differences was computed as was done previously for the speed groups. It was found that the span was significantly longer in the alphabetic sentence test than in the first easy material test, the disconnected word test, or the difficult material test. Apparently the difference was insignificant in all other cases.

In the test made up of disconnected words, some wide differences are found between the spans in writing words of different lengths, as is

evident in Table VII. However, the number of cases was so small that the probable errors become very large. The greatest difference, that between three- and seven-letter words, was 1.66 ± .47. This difference is 3.53 times the probable error, which suggests that the difference may be of some importance. However, this difference and all others in this group may be due entirely to sampling errors.

The distributions of the distances between the eye and the hand, in terms of words, for the various tests, are presented in Table VIII. The indication is that on the tests of about equal difficulty (I, II and V), the spans are approximately the same. It is interesting to note that the use of phrases of four words, with the eye three words ahead of the hands, is not found in the test made up of alphabetic sentences. In Test I the four-word phrases are much the same as those found among the slower writers: "designed but it was"; "for he had daubs"; "was not a good"; "mixed up with the." Test V, although of apparently the same degree of difficulty as Test I, was found not to lend itself quite so readily to the use of phrases expressing a unit of thought. Only two such were found in the record, namely: "he goes on the"; and "willing to pay him." Nevertheless, in three other cases phrases of this length were found: "demand and the better"; "thing well is always"; and "it, know it all."

TABLE VIII.—DISTRIBUTION OF THE NUMBER OF WORDS SEPARATING THE EYE-FIXATION AND THE HAND ON THE VARIOUS TESTS TAKEN BY THE FORMER WORLD'S CHAMPION

Length of span in words	Test number									
	I	V	II	IV	III	(3)	(4)	(5)	(6)	(7)
3	4	5	.....	.....	2	2				
2	13	14	18	6	13	6	5	2		
1	13	19	19	14	36	2	5	13	10	6
0	1	4	.....	15	5	.....	.....	.....	.....	5
Mean.....	1.68	1.48	1.49	.74	1.21	2.00	1.50	1.13	1.00	.55

Among the lines of disconnected words only two cases were found in which the eye was three words ahead of the hand, and those were among the three letter words: "jet our van bad"; and "dig rub out fin." However, the presence of these two groups of such length, and of thirteen examples of a span of two words, gives some evidence that

the eye will occasionally take in a large span ahead of the hands even if there is no thought unit involved. This throws some doubt on the theory that the writing is done by phrases. An equally plausible explanation is that the eye simply moves ahead occasionally and takes in a longer span, irrespective of the thought content. Much more evidence is necessary before a definite conclusion can be reached on this point. It should be remembered, also, that data for the test on disconnected words were secured only for the expert typist. It is possible that the same results might not have been obtained for slower writers.

Conclusions

1. While the variations are not consistent or regular, it may be said that, in general: Rapid writers require fewer fixations per line than do slower writers; the eye-hand span is larger in the case of rapid writers; the more rapid writers carry the span over the end of the line in a larger percentage of the cases.

2. When reading copy for typewriting the subjects make many more fixations per line, and longer fixations, than they do in ordinary reading.

3. The average length of the eye-hand span varies for individual writers from 3.24 spaces to 7.60 spaces, and for speed groups from 3.70 spaces to 6.83 spaces.

4. The shortest span found was zero spaces—the eye resting on the letter that was being struck. The longest span found was one in which the eye was thirteen spaces ahead of the hand.

5. The evidence is that in general an increase in speed of ten words per minute will be accompanied by a statistically significant increase in the average length of the eye-hand span.

6. There is evidently a tendency for the eye to keep an average distance ahead of the hand such that the time interval between seeing the letter and writing it is approximately one second, irrespective of the speed of the writing.

7. In isolated cases, the eye was over 2.5 seconds ahead of the hand.

8. The general tendency, among all but the most rapid writers, is to keep the eye one word ahead of the hand, on the simple type of material used in the tests. Among the most rapid writers the eye may be, on the average, about 1.5 words ahead of the hand on this type of material.

9. In a few isolated instances the eye was found to be three, and in one case four, words ahead of the hand.

10. When the eye is more than one word ahead of the hand it is found that the words involved are, in most cases, those which make up a phrase conveying a unit of thought. However, some exceptions are found.

The following conclusions may be drawn from the results of the tests on different types of material, in the case of the expert typist:

11. When different types of material are used, only slight variations are found in the eye movements themselves.

12. The eye-hand span was found to be significantly longer in the case of a test made up of alphabetic sentences than in one easy selection, a test made up of disconnected words, and a test of difficult material. In all other cases the difference was statistically insignificant.

13. On account of the small number of cases no significant differences could be found among the eye-hand spans on disconnected words of various lengths. Those for five- and seven-letter words were noticeably shorter than those for three-, four- and six-letter words.

14. Cases in which the eye was two and three words ahead of the hand occurred even in the test on disconnected words, indicating that such groupings may be due to eye habits, rather than to writing by phrases conveying a unit of thought.

A COMPARISON OF THE RESULTS OBTAINED FROM THE TEACHING OF SHORTHAND BY THE WORD UNIT METHOD AND THE SENTENCE UNIT METHOD

MILDRED CLARK

Detroit High School of Commerce

AND

D. A. WORCESTER

University of Nebraska

Reading by the sentence unit method was suggested by Comenius as early as 1631. Experiments have been made in reading, penmanship, and music which indicate that there is more rapid progress when a larger unit is used as the basis. In the teaching of shorthand, however, the prevalent method is still to first present principles and to apply these principles by writing isolated words. It seems, therefore, pertinent to inquire if this method is the better. The object of this study, then, is to determine if possible whether or not elementary shorthand can best be learned by beginning with isolated words and gradually going from these to phrases and finally to sentences or by the more comprehensive method of beginning by writing sentences.

The aim of the word unit method—the conventional method—is to teach the principles of shorthand to students so that they may apply them accurately to the words in the text and to a few other common words. At the first meeting of the class the shorthand characters to be taught are written upon the blackboard and the students are taught to recognize them by repeated reading. After careful analysis of the forms there is a writing drill of isolated character strokes. When the vowels are introduced the students are drilled on the names of the groups and the three sounds of each group. On following days words using the new sound or principle are written on the board and students are drilled in reading and writing them. The usual test given is a list of words applying the principles learned.

The aim of the sentence unit method[1] is to familiarize students immediately with the activity as it is going to be practiced and instead of presenting character for character, the word and sentence are

[1] The method employed was that of Miss Gertrude Beers of the University of Nebraska.

presented the first day. It is desired that the student be able to read rapidly and accurately and to write a character for every word dictated, to keep his pencil moving, to recognize good shorthand and to improve his own outlines as rapidly as possible. The material is presented to the class by placing sentences upon the blackboard. At first the teacher must read to the students. Then they read in concert. Finally individual students can read. Words of high frequency are presented and are used over and over again, not in the same sentence but in other and similar sentences. Students can read a new sentence by themselves quite early in the presentation. New sentences are rapidly introduced. Frequent tests are given in reading. They are also tested in their ability to take and transcribe dictation rapidly and accurately.

Procedure of This Investigation

The shorthand class in most high schools is too small to secure reliable results from experimentation; therefore, the cooperation of eleven schools having well-recommended teachers was secured to assist in this experiment. Six of the schools, with one hundred nine students, followed the sentence unit plan of learning and five of the schools, with eighty-three students, followed the word unit plan. None of these students had had any previous training in shorthand. Only two of the teachers had had training in the sentence method. Both of these were teachers of first year experience and they both followed the sentence unit plan. All the other teachers had been trained in the conventional method. The eleven schools were from Montana (1), Illinois (2), Florida (1), Michigan (1), and Nebraska (6). Five tests were given to all students at stated intervals. Because of unavoidable conditions not every student took every test. The Terman Group Test of Mental Ability was given also to every student.

The Tests

Each of the five tests was composed of four parts. The first part of each test was a list of words dictated slowly to be written in shorthand by the student. This is similar to the kind of test given by the conventional method. The second part of each test was dictation material to be dictated at sixty words a minute for the first and second tests, at seventy words a minute for the third and fourth tests, and at

TABLE I.—A COMPARISON OF THE WORD AND SENTENCE METHODS ON FIRST SEMESTER TEST FOR THE ENTIRE GROUP

Part I. Words			Part II. Dictation			Part III. Reading			Part IV. Transcription		
Scores	Word method	Sentence method	Scores	Word method	Sentence method	Scores	Word method	Sentence method	Scores	Word method	Sentence method
47–50	1	10	82–86	2	12	121–125	2	13	171–180	2	18
43–46	2	13	77–81	1	15	116–120	1	19	161–170	4	18
39–42	13	22	72–76	6	16	111–115	2	8	151–160	10	14
35–38	13	23	67–71	7	7	106–110	3	8	141–150	3	10
31–34	14	10	62–66	12	14	101–105	4	11	131–140	8	6
27–30	16	9	57–61	5	10	96–100	3	11	121–130	7	8
23–26	9	10	52–56	3	11	91– 95	3	5	111–120	5	10
19–22	4	1	47–51	9	5	86– 90	5	6	101–110	9	11
15–18	3	5	42–46	6	6	81– 85	6	8	91–100	14	5
11–14	2	4	37–41	10	2	76– 80	7	5	81– 90	5	2
7–10	1	1	32–36	6	3	71– 75	8	4	71– 80	5	1
			27–31	7	1	66– 70	6	0	61– 70	3	0
			22–26	3	3	61– 65	5	1	51– 60	1	2
			17–21	1	3	56– 60	0	1	41– 50	0	2
						51– 55	7	0	31– 40	1	1
						46– 50	4	4	21– 30	1	0
						41– 45	3	0			
						36– 40	3	0			
						31– 35	2	1			
						26– 30	1	1			
						21– 25	3	1			
						16– 20	0	0			
						11– 15	0	0			
						6– 10	0	0			
						1– 5	0	1			
	N 78	*N* 108		*N* 78	*N* 108		*N* 78	*N* 108		*N* 78	*N* 108
Mean	31.72	35.52	.....	51.68	63.34	.....	73.06	96.27	.....	115.24	138.13
SD_{dis}	8.01	22.93	.....	16.35	16.85	.....	24.65	24	.....	32.8	33.5
SD_{mean}	.90	2.2	.....	1.85	1.62	.....	2.79	2.3	.....	3.71	3.22
SD_{diff}	2.3		.....	2.45		.....	3.61		.....	2.03	
EC	.59		.....	1.7		.....	2.3		.....	4.00	

TABLE II.—A COMPARISON OF THE WORD AND SENTENCE METHODS ON SECOND SEMESTER TEST FOR THE ENTIRE GROUP

TABLE II.—A COMPARISON OF THE WORD AND SENTENCE METHODS ON SECOND SEMESTER TEST FOR THE ENTIRE GROUP

Part I. Words			Part II. Dictation			Part III. Reading			Part IV. Transcription		
Scores	Word method	Sentence method	Scores	Word method	Sentence method	Scores	Word method	Sentence method	Scores	Word method	Sentence method
24–25	7	7	151–160	2	1	96–100	7	45	151–160	2	19
22–23	4	22	141–150	0	7	91– 95	10	22	141–150	5	7
20–21	4	18	131–140	1	11	86– 90	7	4	131–140	10	20
18–19	6	12	121–130	2	14	81– 85	6	5	121–130	4	7
16–17	4	7	111–120	12	15	76– 80	4	0	111–120	9	10
14–15	1	6	101–110	8	17	71– 75	4	1	101–110	5	8
12–13	7	4	91–100	15	8	66– 70	6	1	91–100	4	6
10–11	8	3	81– 90	3	5	61– 65	1	3	81– 90	6	5
8– 9	7	4	71– 80	4	4	56– 60	1	2	71– 80	4	0
6– 7	6	1	61– 70	5	2	51– 55	2	1	61– 70	2	2
4– 5	0	2	51– 60	4	2	46– 50	2	0	51– 60	2	1
2– 3	2	1	41– 50	0	2	41– 45	0	2	41– 50	2	0
0– 1	0	1				36– 40	1	1	31– 40	0	2
						31– 35	1	0	21– 30	0	0
						26– 30	0	1	11– 20	0	1
						21– 25	2	0	1– 10	1	0
						16– 20	1	0			
						11– 15	0	0			
						6– 10	1	0			
	N 56	*N* 88		*N* 56	*N* 88		*N* 56	*N* 88		*N* 56	*N* 88
Mean.....	14.97	18.61	.....	98.14	111.57	.....	75.65	89.97	.....	107.6	123.61
SD_{dis}	6.41	5.48	.....	22.65	23.78	.....	27.38	15.33	.....	31.67	30.04
SD_{mean}.....	.85	.58	.....	3.02	2.53	.....	3.65	1.63	.....	4.23	3.2
SD_{diff}.....	1.02		.....	3.93		.....	3.99		.....	5.3	
EC.....	1.2		.....	1.2		.....	1.2		.....	1.0	

eighty words a minute for the fifth test. The third part was a group of shorthand sentences to be written in longhand by the student. The fourth part was transcription of shorthand notes which the student had written from dictation at the same speed as the dictation given in Part II. Dr. Horn's list of one thousand most commonly used words was used in constructing the tests. Definite uniform instructions were sent to each teacher for guidance in giving the tests.

Results Obtained

There will be no attempt here to present all of the data but the complete distribution for the test given at the close of the first semester and for the test given at the close of the second semester will be shown together with a summary of the scores on the other tests. Table I shows results obtained from the entire group of students in test 3 given at the close of the first semester.

It will be seen that with the exception of Part I there is high statistical significance in favor of the sentence group. Even in Part I the results favor this group although the subject-matter of this portion of the test is that typically employed by the word method.

Table II shows results obtained from the entire group of students in test 5 given at the close of the second semester.

The results here are similar to that found in test 3. Even in Part I of the test, the word list, the sentence group is far superior. In every case the results are statistically significant and in favor of the sentence group.

Results were also obtained from pairs of students who were matched according to raw scores on the Terman tests. When there were two students in one group who had the same raw score and one in the other group the student in the first group who had a chronological age nearest the one in the second group was chosen to complete the pair. When there were two with the same raw score in each group they were paired according to nearest chronological age. Since all of the pupils in this investigation were either in the eleventh or twelfth grade of high school it seemed unnecessary to employ any more involved procedure for the matching. The results obtained from the pairs for tests 3 and 5 are shown in Tables III and IV.

Here again the sentence method proves to be far superior to the word method. In every case the comparison is in favor of the sentence unit plan.

TABLE III.—A COMPARISON OF THE WORD AND SENTENCE METHODS ON FIRST SEMESTER TESTS FOR THE PAIRED GROUP

Part I. Words			Part II. Dictation			Part III. Reading			Part IV. Transcription		
Scores	Word method	Sentence method	Scores	Word method	Sentence method	Scores	Word method	Sentence method	Scores	Word method	Sentence method
48–50	0	4	81–85	1	7	121–130	1	6	171–180	0	8
45–47	2	1	76–80	0	6	111–120	1	10	161–170	2	6
42–44	0	7	71–75	5	5	101–110	6	7	151–160	6	7
39–41	3	3	66–70	4	5	91–100	2	6	141–150	1	4
36–38	4	10	61–65	3	2	81– 90	4	6	131–140	8	1
33–35	7	7	56–60	5	5	71– 80	11	6	121–130	4	2
30–32	4	3	51–55	2	6	61– 70	5	0	111–120	2	5
27–29	10	1	46–50	4	1	51– 60	3	1	101–110	3	6
24–26	4	3	41–45	3	1	41– 50	4	1	91–100	9	3
21–23	4	1	36–40	7	1	31– 40	3	0	81– 90	4	1
18–20	2	0	31–35	5	3	21– 30	4	1	71– 80	3	0
15–17	1	3	26–30	3	0				61– 70	1	0
12–14	1	0	21–25	1	1				51– 60	0	0
9–11	2	1	16–20	1	1				41– 50	0	0
									31– 40	0	1
									21– 30	1	0
	N 44	*N* 44		*N* 44	*N* 44		*N* 44	*N* 44		*N* 44	*N* 44
Mean	29.52	35.66	.....	50.09	63.16	.....	71.46	98.5	.....	115.55	138.5
$SD_{dis.}$	8.25	9.21	.....	16.3	17.2	.....	25.58	22.27	.....	30.89	32.34
SD_{mean}	1.24	1.38	.....	2.46	2.59	.....	3.85	3.35	.....	4.65	4.87
$SD_{diff.}$	1.85		.....	3.57		.....	5.1		.....	6.73	
EC	1.1		.....	1.3		.....	1.9		.....	1.2	

TABLE IV.—A COMPARISON OF THE WORD AND SENTENCE METHODS ON SECOND SEMESTER TEST FOR THE PAIRED GROUP

Part I. Words			Part II. Dictation			Part III. Reading			Part IV. Transcription		
Scores	Word method	Sentence method	Scores	Word method	Sentence method	Scores	Word method	Sentence method	Scores	Word method	Sentence method
24–25	4	1	151–155	1	0	96–100	2	14	151–160	1	5
22–23	2	8	146–150	0	0	91– 95	8	7	141–150	1	1
20–21	1	6	141–145	0	1	86– 90	1	0	131–140	5	7
18–19	2	3	136–140	0	2	81– 85	2	2	121–130	1	3
16–17	2	4	131–135	0	3	76– 80	3	0	111–120	4	4
14–15	0	1	126–130	0	2	71– 75	0	1	101–110	3	2
12–13	3	0	121–125	0	2	66– 70	4	1	91–100	1	1
10–11	5	1	116–120	1	3	61– 65	1	0	81– 90	3	1
8– 9	3	1	111–115	4	2	56– 60	0	0	71– 80	2	0
6– 7	4	0	106–110	1	2	51– 55	1	1	61– 70	2	1
4– 5	0	1	101–105	2	5	46– 50	1	0	51– 60	2	1
2– 3	1	0	96–100	2	1	41– 45	0	1	41– 50	1	0
0– 1	0	1	91– 95	6	1	36– 40	1	0	31– 40	0	1
			86– 90	3	0	31– 35	0	0	21– 30	0	0
			81– 85	0	0	26– 30	0	0	11– 20	0	0
			76– 80	0	0	21– 25	1	0	1– 10	1	0
			71– 75	1	0	16– 20	1	0			
			66– 70	3	0	11– 15	0	0			
			61– 65	0	0	6– 10	1	0			
			56– 60	3	1						
			51– 55	0	1						
			46– 50	0	1						
	N 27	*N* 27		*N* 27	*N* 27		*N* 27	*N* 27		*N* 27	*N* 27
Mean	14.49	18.56	.....	93.32	110.91	.....	72.76	90.35	.....	99.71	121.18
$SD_{dis.}$	6.64	5.76	.....	21.3	24.25	.....	24.75	14.15	.....	35.08	30.59
SD_{mean}	1.27	1.1	.....	4.09	4.66	.....	4.76	2.72	.....	6.75	5.88
$SD_{diff.}$	1.68		.....	6.2		.....	5.48		.....	8.95	
EC	.8		.....	1		.....	1.1		.....	.8	

Table V shows in summary the results of tests 1, 2, and 4 for both the entire group and the paired group.

In every instance save one, in which an error was made in the instruction of one school, the results of these tests are also conclusively in favor of the sentence group.

Looked at from another point of view it was noted that out of six hundred ninety chances the paired pupils of the sentence group were superior four hundred ninety-one times and equal to the word group fourteen times. In other words, the paired students in the sentence group equalled or surpassed the word group in seventy-three per cent of the cases.

TABLE V.—A SUMMARY OF TESTS 1, 2, AND 4 FOR THE ENTIRE GROUP

Test 1

	Part I		Part II		Part III		Part IV[1]	
	W[2]	S[3]	W	S	W	S	W	S
Number	51	108	51	108	51	108	51	108
Mean score	23.45	24.027	68.21	72.2	78.11	90.63	58.79	69.65
$SD_{dis.}$	4.26	4.6	19.51	16.77	19.55	12.25	38.63	22.552
SD_{mean}	.59	.44	2.73	1.61	2.73	1.17	1.48	2.17
$SD_{diff.}$	.73		3.17		2.97		2.62	
EC	.28		.45		1.5		1.4	

For the Paired Group

Number	28	28	28	28	28	28	22	22
Mean score	23.43	24.93	66.72	70.12	80.1	88.85	73.5	72.55
$SD_{dis.}$	4.16	5.08	18.9	20.65	18.3	17.7	26.5	23.95
SD_{mean}	.78	.96	2.41	3.9	3.45	3.34	5.65	5.1
$SD_{diff.}$	1.23		4.58		4.8		7.61	
EC	.43		.26		.6		.04	

[1] An error was made in giving instructions in one school so that these scores were omitted in the paired group.

[2] "W" here stands for the word method group.

[3] "S" here stands for the sentence method group.

TABLE V.—*Continued*

Test 2

	Part I		Part II		Part III		Part IV	
	W	S	W	S	W	S	W	S
Number	83	103	83	103	83	103	83	103
Mean score	32.14	38.74	42.09	50.81	35.71	47.845	98.89	137.45
$SD_{dis.}$	9.12	7.64	12.2	8.92	18.02	15.04	41.8	33.9
SD_{mean}	1	.75	1.33	.88	1.97	1.48	4.58	3.34
$SD_{diff.}$	1.4		1.59		2.46		5.66	
EC	1.6		1.9		2		2.4	

For the Paired Group

Number	43	43	43	43	43	43	43	43
Mean score	31.127	38.8	40.152	50.058	34.22	50.24	95.535	138.558
$SD_{dis.}$	8.13	7.86	11.46	8.31	17.67	12.42	43.07	32.92
SD_{mean}	1.24	1.19	1.74	1.26	2.69	1.89	6.56	5.02
$SD_{diff.}$	1.71		2.14		3.28		8.26	
EC	1.6		1.6		1.7		1.8	

Test 4

	Part I		Part II		Part III		Part IV	
	W	S	W	S	W	S	W	S
Number	56	97	56	97	56	97	56	97
Mean score	18.08	21.33	91.72	107.44	70.03	79.45	106.36	125.79
$SD_{dis.}$	6.714	6.12	20.07	22.47	14.55	10.695	21.7	23.72
SD_{mean}	.88	.62	2.68	2.28	1.94	1.08	2.89	2.4
$SD_{diff.}$	1.07		3.51		2.22		3.75	
EC	1.0		1.6		1.5		1.8	

For the Paired Group

Number	32	32	32	32	32	32	32	32
Mean score	16.5	20.69	87.57	106.31	68.344	79.125	101.94	115.69
$SD_{dis.}$	6.54	6.08	19.8	22.7	15.1	9.2	22.34	22.42
SD_{mean}	1.15	1.07	3.5	4.01	1.47	.89	3.94	3.96
$SD_{diff.}$	1.57		5.32		1.71		5.58	
EC	.9		1.2		2.2		.8	

Conclusions

In this investigation of relative merits of the word unit method and the sentence unit method of teaching shorthand in which five schools employed the word unit method and six the sentence unit method, it appears that in every test the sentence unit method proved superior. This was in spite of the fact that four out of the six teachers using the sentence method had been trained only in the word method. This was true both for the entire group and for the paired group. In a few cases the difference was not statistically significant but the variation was in the same direction as those that were statistically significant. The study shows clearly the superiority of the sentence unit method.

It should perhaps be noted that students seem to find it somewhat more difficult to grasp the presentation in the sentence unit method the first day or two, there is more concentration required, and it takes more initial effort to learn shorthand by the sentence unit method, but there is much time gained through the process since the "word" habit is not formed, later to be broken up and another habit formed to take its place. The individual has more confidence in himself and is more interested in the sentence unit method than in the word unit method. The sentence unit plan brings out to the learner the fact that he must learn rules—not verbatim—but by the application of them. The method exemplifies the rule that one should always begin by doing a thing as nearly as possible in the way it is eventually to be done.

THE RETENTION OF CERTAIN LATIN SYNTACTICAL PRINCIPLES BY FIRST AND SECOND YEAR LATIN STUDENTS AFTER VARIOUS TIME INTERVALS[1]

LEO R. KENNEDY

The Creighton University

Previous Work

Although a monumental amount of experimental work has been done in the field of learning and retention since Ebbinghaus derived his curve of forgetting, yet it has been only in relatively recent years that investigators have directed their attention to a quantitative determination of classroom learning and retention. A review of the literature relating to this topic reveals fifteen studies in the elementary school field, the most comprehensive of which are those of Irmina[2] and Noonan.[3] Nine investigations have been reported which involve either high school or college students. However, six of these do not afford us any valid measure of the remembrance of subject matter because the initial achievement of the students was not determined. The studies of Johnson,[4] Schmidt,[5] and Worcester[6] are the only ones on this level that determine remembrance in terms of initially possessed information. Of these three that of Johnson is the most comprehensive and most recent. The subjects of this study were one hundred twenty-eight college Botany students. At the close of each quarter during the academic year of 1926–1927 a different objective examina-

[1] The complete study is on file in the office of Dr. D. A. Worcester, Teachers College, University of Nebraska. The author is deeply obligated to Dr. Worcester for aid in conducting this research.

[2] Irmina, Sister: "The Effects of Summer Vacation upon Retention of the Elementary School Subjects." Catholic University of American Press, Washington, D. C., pp. 99, 1928.

[3] Noonan, M.: "The Influence of the Summer Vacation on the Abilities of Fifth and Sixth Grade Children." Teachers College, Columbia University, pp. 103, 1926.

[4] Johnson, P.: The Permanence of Learning in Elementary Botany. *Journal of Educational Psychology*, Vol. XXI, 1930, pp. 37–47.

[5] Schmidt, A.: "The Effect of Objective Presentation on the Learning and Retention of a Latin Vocabulary." Loyola University Press, Chicago, pp. 192. 1923.

[6] Worcester, D. A.: The Permanence of Learning in High School Subjects—Algebra. *Journal of Educational Psychology*, Vol. XIX, 1928, pp. 343–344.

tion was given these students. In all there were two hundred sixty-five true-false and three hundred twenty-two completion questions. On the first day of school in the fall a selected list of questions from each of the three former examinations was given the subjects. The study is primarily concerned with a group of forty-two students who had taken the previous examinations. For twenty-four subjects the percentage of loss in retention was 43.3 per cent over the three quarters combined. For a six and nine month interval combined, thirty-six students lost 47.8 per cent. Those who had not studied Botany for fifteen months showed a loss of 74.1 per cent. There was a deterioration of 76.1 per cent for those who had no Botany for twenty-seven months.

The materials in Schmidt's study were thirty Latin nouns, thirty adjectives, and forty verbs. The subjects were drawn from grades six and eight. One-half of the words were learned objectively and one-half non-objectively. Six periods were devoted to the learning of fifty words. At the close of the last period a fifteen minute recall test was given. In this test the students were required to give the English equivalents of the Latin words. The same procedure was followed for the second group of fifty words. Four weeks later a retest was given. The percentages of retention at this time were as follows: Seventy-six per cent retention of words learned by the objective method and seventy-eight per cent retention of materials learned by the non-objective method. Three months after their first and only previous recall the percentages of remembrance for one hundred thirty-three of these pupils was fifty-seven per cent for words learned by the objective method and sixty per cent for words learned by the non-objective method.

Worcester has measured the degree of retention in Algebra by means of the Douglas Algebra Test. In February Form A-1 was given to a group of students; about three weeks later the same subjects took Form A-2 of this test, and midway in May Form B-1 was administered to them. After an interval of seven months these tests were repeated. Sixteen of the students equalled or surpassed their original scores on Form A-1. On Form A-2 only five subjects equalled or exceeded their initial scores on this test. The percentage of retention on Form A-2 to the original score was eighty-four. The per cent of remembrance on A-1 to the initial achievement in A-2 was eighty. The author suggested that these high percentages of retention were due in part, at least, to over-learning. Evidence to support this contention

was found in the retention percentage for Form B-1, which was found to be thirty-three. The subject-matter covered in this particular test form comes toward the close of the school year, thereby limiting the opportunity for over-learning.

The Problem of This Study

Some of the more important factors that we have attempted to describe in quantitative terms in this study of Latin syntax are:

1. The amount of remembrance of total initially correct and incorrect responses at stated time intervals.
2. The amount of retention of identically correct and identically incorrect[1] responses at the given time intervals.
3. The relation of general intelligence to the retention of Latin syntax as measured by the method of recognition.
4. The effect of study during September on the regaining of pre-vacation achievement.
5. The influence of modern language study on the remembrance of Latin syntax.
6. Grade differences, inter-group and intra-group differences; sex differences.
7. Types of syntactical principles best remembered.

The general plan of the investigation was as follows: All the Latin 2 students (those in the second semester of first year Latin) and all the Latin 4 students (those in the second semester of second year Latin) of the Lincoln, Nebraska, Public Schools were given the Pressey Latin Syntax test[2] at the close of the school year in 1929. On the first day of school in September all of those who had taken the pre-vacation test, and who were continuing with Latin, were retested with the same form[3] of the Pressey test. Hereinafter these subjects are termed *continuation group students*. On the same day all other students who had taken the pre-vacation test and who were registered for the fall term of school, but who were not continuing Latin, were also retested. In this study these subjects are termed *stop group students*. The

[1] By identically correct or incorrect response is meant the same response made twice or more in succession by the same individual. The opportunity for studying such responses was present because a multiple choice test was used.

[2] The Pressey test is of the four response multiple-choice type. It illustrates thirty-two principles of Latin syntax.

[3] There is only one form of this test; however, a thorough study was made of practice effect, using students other than those in this investigation, and this factor was found to be statistically insignificant.

continuation group students of both Latin 2 and 4 were retested for the third time one month after the resumption of school, in order to determine the influence of study on the regaining of initial achievement. The stop groups were given their third test in May 1930. The population of the groups involved in this study follow: Latin 2 continuation group students one hundred fifteen, Latin 4 continuation group twenty-two, Latin 2 stop group fourty-four, and Latin 4 stop group sixty-three. All of the above students participated in all of the tests proper to their respective groups. None of the students had studied Latin during the summer vacation.

The Retention of Correct Responses

In order to determine the amount of Latin syntax actually retained by the various groups the percentages of remembrance in terms of initial scores were derived for the various time intervals. In all groups the percentages of remembrance for identically correct responses over the summer vacation exceed that which chance[1] would yield, thus indicating a certain permanence of learning. These percentages, which are in excess of the chance percentages, are: For the Latin 2 continuation group two per cent, Latin 2 stop group six per cent, Latin 4 continuation group fifty-one per cent, and the Latin 4 stop group twenty-seven per cent. The large differences in the percentages of retention between the continuation and stop groups constitute an interesting and instructive piece of information. It appears that intention to continue with the study of Latin is a very important factor in terms of the amount of initial knowledge retained over the summer vacation. The importance of this element is further appreciated when it is found that the reliability of the difference[2] between the mean initial knowledge of the Latin 2 stop and continuation groups and the Latin 4 stop and continuation groups is 2.08 and .02 respectively, which indicates statistical insignificance. The analysis of other data collected in this investigation tend to corroborate the above inference.

The best measure of retention of identically correct responses over a long period of time, namely one year, is offered us by the two stop

[1] Since the Pressey test is of the multiple type sort it is obvious that a certain amount of guessing will occur. In view of this condition obtained percentages are considered significant only if they exceed that which chance would allow. It is recognized that the criterion of chance is not wholly valid in this connection, but it is the most adequate reference point we as yet possess.

[2] See Garret: "Statistics in Psychology and Education." Pp. 118–136. This was the measure of significance used throughout this study.

groups. The Latin 2 stop group retains its initial knowledge six per cent in excess of chance and the Latin 4 stop group remembers twenty per cent in excess of what chance would allow. The percentage for the Latin 4 group may be considered very high in view of the fact that intention has been shown to be an important factor in retention.

A more liberal measure of remembrance may be had by comparing the total amount known in September to the total amount known before vacation. In terms of this criterion the percentages of remembrance for the various groups are: Latin 2 continuation group seventy; Latin 2 stop group seventy-one; Latin 4 continuation group eighty-five; and the Latin 4 stop group sixty-six. The percentages derived on this basis are higher than those reported in similar studies, and although some of the magnitude of the percentages is explainable in terms of the type of test used, we are still justified in inferring that students do remember considerably more of their subject-matter than that with which they are commonly credited. When we apply the same criterion as above stated to the interval of one year, we find the Latin 2 stop group retains sixty-eight per cent of their syntactical information and the per cent for the Latin 4 stop group becomes fifty-four. It is suggested that this difference of fourteen per cent in favor of the Latin 2 group is due to the fact that this group was more recently in contact with the principles exemplified in the Pressey test.

The Retention of Error

The analysis of the identical error situation presents a challenging set of facts. The data show that all groups remembered identical errors over the summer vacation in excess of the chance percentage. The Latin 2 continuation group retained eighteen per cent of their initial identical errors, the Latin 2 stop group eight per cent, the Latin 4 continuation group nineteen per cent and the Latin 4 stop group five per cent. The above percentages are in excess of what chance would allow. No wholly adequate reason is known for the greater retention of identical errors over the summer vacation on part of the continuation groups, but it is suggested that since these two groups initially knew more correct responses they also retained more identically incorrect responses. It is also possible that in this instance intention has again been a factor, but operating detrimentally to the student.

In terms of the chance percentage, with the test used, we would expect a 6.75 per cent retention of identical errors over three test periods, which in this instance covers a period of one year. However,

the actual percentages, in excess of chance, are for the Latin 2 stop group 8.75 and for the Latin 4 stop group 11.75. The percentages on this basis for the Latin 2 and 4 continuation groups are 12.75 and 11.75 respectively. These latter two groups received their third test one month after the resumption of school in the fall. These percentages indicate that one month of study in the fall has been somewhat effective in reducing the identical error configuration for the Latin 2 continuation group. This may be seen if we compare the 12.75 per cent with the eighteen per cent. A similar comparison of the Latin 4 continuation group leads us to the same inference. In this connection it might be noted that the analysis of the data also indicated that one month of study in the fall was about three times more effective in breaking down identical errors made in May and September than in adding to the identically correct responses made in the same period. *The whole identical error situation indicates that all groups of students were actually learning and remembering erroneous responses.* This only further emphasizes the great necessity of obtaining correct initial response configurations.

The Effect of Intelligence

In order to throw some light on the relation between initial achievement and retention of syntax and the factor of intelligence, zero order and partial coefficients of correlation[1] were derived. A few of these are given here. The probable errors are omitted because they are all very small. The zero order correlations between total amount known before and after summer vacation are for the Latin 2 continuation group .76, the Latin 2 stop group .74, the Latin 4 continuation group .90, and for the Latin 4 stop group .74. The partial correlations, holding intelligence constant, for this same relationship are .75, .75, .89, and .76 respectively, thus indicating that, in general, intelligence[2] is not a significant factor in the remembrance of Latin syntax over the summer vacation. Nor does intelligence bear a high relation to the amount of Latin syntax initially learned by the groups as a whole, because the zero order correlations between intelligence and initial achievement in no group exceed .25 and range as low as .01.

[1] These are product-moment coefficients.

[2] It should be remembered, however, that these students do represent a selection in terms of intelligence. The mean intelligence quotients for the Latin 2 continuation group, Latin 2 stop group, Latin 4 continuation group, and Latin 4 stop group are respectively 110, 106, 115, and 109.

The zero correlations between identically correct responses in the May 1929 and September tests yield coefficients of .85 for the Latin 2 continuation group, .74 for the Latin 2 stop group, .87 for the Latin 4 continuation group, and .84 for the Latin 4 stop group. When intelligence is partialled out these relationships become .92, .75, .94, and .85 respectively.

The zero coefficients for the Latin 2 and Latin 4 continuation groups between total amount known in May 1929 and October—that is after one month of study in the fall—are .70 and .87 respectively. These are practically the same as those found for these groups over the summer vacation. The partial coefficients of .66 and .87 again indicate that intelligence, in general, is not a factor in retention of Latin syntax. It is interesting to note the correlations among stop groups between total amount known in May 1929 and total amount known in May 1930. For the Latin 2 stop group the coefficient is .12 and for the Latin 4 stop group it is .23. These figures are not large enough to be of definitely useful prognostic value. Here again the partialling out of intelligence does not affect the magnitude of the coefficients.

The zero order correlations between total errors in May 1929 and total errors in September for the various groups follow: Latin 2 continuation group .72, Latin 2 stop group, .83, Latin 4 continuation group .90, and Latin 4 stop group .92. When intelligence is held constant these figures become respectively .64, .84, .89, and .89. This situation may be interpreted to mean that intelligence, although showing a tendency to be negatively correlated with error, does not exert any significant influence. The coefficients for the stop groups between total errors in May 1929 and total errors in May 1930 are .81 for the Latin 2 stop group and .84 for the Latin 4 stop group. The partial coefficients are .83 and .81 respectively. It will be noted the zero coefficients are of about the same magnitude as those for the same two groups over the summer vacation. An interesting aspect of the problem is presented in the relationship between identical errors in May 1929 and September. For the Latin 2 continuation group the coefficient is .48, for the Latin 2 stop group .28, for the Latin 4 continuation group .91, and for the corresponding stop group .87. When intelligence is held constant the relationships are respectively .47, .28, .90, and .86. It is seen that the coefficients are much lower for the two Latin 2 groups than for the Latin 4 groups. This situation might be taken to mean that one year of Latin is not sufficient to fix well in the minds of the pupils their remembrance of identical errors,

whereas the magnitude of the correlations for the Latin 4 groups would indicate that these students have learned their errors well.

The general inference to be made from the correlations given in this section seem to be that, for these groups as a whole, intelligence is not a significant factor in the acquisition of Latin syntax nor in its retention. Nor is it distinctly helpful in the avoidance of error. Achievement is the significant element. Those who know most Latin syntax will remember most, both in terms of correct and incorrect responses.

Differences in Retention between the Upper and Lower Quarters in Terms of Intelligence

This part of the investigation was undertaken with the purpose of shedding light on the question of absolute and relative retention. This problem is often raised in discussion, but much neglected in studies of remembrance of subject-matter. The question may be put: Do those who are of high intelligence[1] learn and retain more, either from an absolute or relative point of view, than those who are of low intelligence? Obviously the answer is important to education. Because if those who are less intelligent forget the same absolute amount as those who are very intelligent, then provision for such a condition should be made in school work. It would mean that in many instances academic work would possess few long-time functional values for the dull student, and particularly so if he stopped attending school. If he continued in his school work it would mean that much drill and review would have to be provided in order to counteract the effects of rapid forgetting.

Because of the magnitude of the population only the Latin 2 continuation group and the Latin 4 stop group were studied. The mean intelligence quotients of the Latin 2 group for the upper and lower quarters was 127 and 95.6[2] respectively with accompanying standard deviations of 7.25 and 6.05. The data indicated that the upper quarter knew a statistically significant greater amount of syntax initially. The upper quarter remembered eighteen per cent more identically correct responses over the vacation period than did the lower division, and from an absolute point of view they remembered

[1] Although the analysis here presented is in terms of intelligence levels the evidence showed that the same inferences are valid when the data are analyzed from the viewpoint of initial achievement.

[2] Intelligence quotients were derived from the Terman Group Test of Mental Ability.

about twice as much. In terms of total amount of correct responses retained over the vacation the upper group remembered three per cent more than the lower quarter. From an absolute viewpoint the upper division remembered a statistically significant greater mean number of items. With regard to errors the upper division committed eight per cent more total errors over vacation than the lower quarter. From an absolute viewpoint the upper group made significantly fewer total errors in September than did the lower division. Relatively, the upper quarter retained five per cent more identical errors over the vacation period than the lower quarter, but absolutely fewer of them, although this difference is not statistically reliable.

The mean intelligence quotient of the upper quarter of the Latin 4 stop group was 125; that of the lower quarter 96. In the Latin 4 stop group the upper quarter initially knew a significantly larger mean amount of syntax. Over the summer vacation the upper group retained 8 per cent more total correct responses than the lower division. On the basis of identically correct responses remembered over the vacation the data showed the upper division remembered thirteen per cent more than the lower quarter, and from an absolute viewpoint about one and three-fourths as many items. Analysis of identically correct responses over the three test periods, which covers an interval of one year, also indicated a thirteen per cent superiority in favor of the upper quarter. That is, an interval of one year's absence from the study of Latin syntax does not cause any change in the difference between the relative percentages of retention made by the upper and lower divisions as compared with the period extending over the summer vacation. Absolutely, the upper group remembered somewhat over two and one-half times as many identically correct responses as the lower quarter over the one year interval. When we consider the total correct responses over a one year period we find the upper quarter to be only one per cent superior to the lower quarter, although on an absolute basis the upper division remembered one-third more items than the lowest group. With regard to errors we find that the upper quarter committed relatively six per cent more total errors after vacation than did the lower group, but absolutely fewer of them. Over a period of one year the upper quarter made forty-one per cent more total errors than the lower group, but absolutely somewhat fewer errors. When we analyze the identical errors over the vacation period we find the upper group retaining six per cent more of these than the lower quarter, but from an absolute viewpoint they retained fewer of them. Over a period of one year the upper and lower

quarters retained relatively an equal amount of identical errors, but absolutely the upper division remembered a few less, although the difference is not statistically reliable.

The general conclusions to be drawn from the analyses of the upper and lower quarters of both the Latin 2 continuation group and the Latin 4 stop group are: (1) Those students who are more intelligent know and retain more Latin syntax, both absolutely and relatively,[1] than those who are less intelligent. (2) Those who are more intelligent also retain a greater relative, but not absolute, amount of their errors. (3) Those who are less intelligent retain, absolutely, more of their errors, but relatively fewer of them than those who are more intelligent.

Effects of One Month of Study in the Fall on the Knowledge of Latin Syntax

In order to determine the influence of study on relearning after the summer vacation the two continuation groups were tested one month after the beginning of school in the fall. The pre-vacation mean of the Latin 2 continuation group was 14.45 and the October mean was 17.37, with a standard deviation of 5.18 in each instance. Hence the month of study was more than equal to the task of returning this group to its initial level of achievement. The reliability of the difference between the means of these two tests was 4.11, which indicates a statistically significant difference. A different situation obtains in the Latin 4 continuation group. The May mean of their group was 23.40 with an SD of 6.05. The October mean was 21.09 with an SD of 6.22. The reliability of the difference is 1.94, which indicates insignificance. This situation is explained on the basis that the Latin 4 continuation group immediately proceeded to the study of Cicero in September without reviewing syntax.

Influence of the Study of Modern Languages on the Retention of Latin Syntax

With the hope of learning something of the influence of modern language study on the remembrance of Latin syntax a comparison was made between all students taking a modern language and all those in the same group who were not following a modern language. Only the Latin 2 and 4 stop groups were employed, because those in

[1] It is to be remembered that this applies to those at the extremes of the population. As has been shown above, intelligence is not a significant factor in the group taken as a whole.

the continuation groups were studying no language other than Latin. The distribution of the students among the modern languages was as follows: In the Latin 2 stop group three students were in French, thirteen in Spanish, and two in German. The distribution for the Latin 4 stop group showed eleven in French, seven in Spanish, and three in German. All students were registered in the first year of the language being studied.

In the Latin 2 stop group the data showed the non-modern language group to be twelve per cent superior to the modern language group in retention of identically correct responses over the summer vacation. It should also be kept in mind that the non-language group had a slightly higher initial mean score. The percentages of remembrance are almost equal for the modern languages and non-language groups in terms of identically correct responses for May 1929, September, and May 1930—that is for the three test periods taken together. Both language and non-language groups remembered exactly the same amount of total correct responses over the summer vacation. However, when we compare the total corrrect responses known in May 1930 with those known in May 1929—that is, after an absence of one year from Latin—we find the modern language group to be twenty-eight per cent superior to the non-language group. This percentage is equivalent to a reliability difference of 3.02, which indicates a significant difference. This might be taken to mean that although the study of a modern language does not aid the student in retaining identically correct responses over a period of one year, its study does help him in a very definite way to retain his total correct amount of Latin syntax when the interval is one year. Further corroboration of this inference was found when the data revealed that those students of the Latin 2 stop group who were studying a modern language scored thirteen per cent more correct responses in the May 1930 test than in the September test. Still additional evidence, although of a negative sort, that the study of a modern language has definitely helped the student to remember his Latin syntax is found among the non-modern language students of the Latin 2 stop group. These students show a decrease of fifteen per cent from September to May 1930 as contrasted with the gain just above given of thirteen per cent in the modern language group, despite the fact that both groups knew exactly the same amount of syntax in September. The mean intelligence quotient of the modern language group is 107, and that of the non-language group 104.5. Therefore intelligence does not explain this situation. Besides it has been shown in a earlier section

of this article that intelligence, in general, is not a significant factor in the acquisition or retention of Latin syntax.

A different situation presented itself when an analysis was made of the modern language and non-language students of the Latin 4 stop group. In every comparison made, these two groups were almost identical. In no instance was there found any beneficial results from the study of modern language in terms of remembrance of Latin syntax. The author is unable to wholly explain why the apparent beneficial effects of modern language study in the Latin 2 stop group did not extend themselves to the Latin 4 stop group. However these suggestions are offered. First, since the Latin 4 students were removed one year from the drill work on syntax that is common in first year Latin, they might not be expected to so well perceive the syntactical similarities of Latin and the modern languages as those who had been more recently subjected to intensive drill in these principles. Another explanation might possibly be found in the fact that most of the language pupils in the Latin 2 stop group were registered in Spanish and that Spanish transferred more than did French, which claimed the majority of modern language pupils in the Latin 4 stop group. Thirdly, it is also possible that the Latin 4 group had already mastered elementary syntax to the degree that schools ordinarily demand. A final consideration is that transfer does not occur in cases where there is relatively high initial learning.

Differences between Grades, Groups, and Sub-groups

The reliability of the difference between the means of various groups and sub-groups was worked out in twenty-one instances. A few of the more significant findings are presented here. The data revealed the Latin 2 continuation group to be somewhat superior to the corresponding stop group in amount of syntactical information in May 1929. However, the continuation group retained significantly more syntax over vacation. The stop group remembered, over vacation, a significantly larger number of identical errors.

The situation is somewhat different in the Latin 4 continuation and stop groups. Each group possessed the same mean amount of initial knowledge. But the continuation group retained a significantly greater mean amount of both total and identically correct responses over the summer vacation. This situation again indicates that intention to continue with the study of Latin is a very important element in keeping intact the learning configurations. The difference

between these groups in the remembrance of identical errors over vacation is not significant. The reliability difference is 1.78.

In all the factors studied the Latin 4 continuation group was significantly superior to both the Latin 2 continuation and stop groups. The Latin 4 stop group is also significantly superior to both the Latin 2 groups.

A very detailed study was made of sex differences, but in no instance was any statistically reliable sex difference found. However, in all cases the girls tended to a superiority over the boys. This is especially true when we compare boys below the group median intelligence quotient with girls below the group median intelligence quotient. Probably this situation is due to greater effort on the part of the girls.

Principles of Syntax Retained

An extensive analysis was made of the thirty-two principles of syntax which are exemplified in the Pressey test. The forms best known initially by both Latin 2 groups were the accusative case, place to which, ablative of accompaniment, and the predicate nominative. Those principles least well known initially were the ablative of respect, ablative of comparison, ablative with certain deponent verbs, ablative absolute, dative of possession, and the dative with certain intransitive verbs. Of these forms the dative of possession was not studied in Latin 2. In general those principles best known initially were also best remembered at the various succeeding test periods. From the viewpoint of retention of identical errors over the summer's vacation and over a period of one year the use of the dative with certain intransitive verbs ranked highest in the Latin 2 groups. Next in rank was the use of the relative pronoun.

In the Latin 4 groups the principles best known initially included all those forms best known initially by the Latin 2 groups, and in addition embraced the accusative with the infinitive, place from which, and the ablative absolute. These principles were also the ones best remembered over the summer vacation. The forms least well retained over the vacation interval were the ablative of comparison, ablative of cause, partitive genitive, predicate adjective, and the dative of possession. After an absence of one year from Latin no principle of syntax was known by as many as two-thirds of the group.

The general conclusions made from the study of the types of syntax found in the test used were: Those principles of syntax best

known initially by all groups also appear to be best known at succeeding test intervals. The opposite of this is also true. It also appeared that only those principles that were known initially by more than two-thirds of a given group had any chance of appearing in the middle or upper one-third of frequency in retests. In general, errors made in May 1929 by from one-third to two-thirds of the students persist on this same frequency level at following test periods. This is particularly true of the stop groups. Principles best known by the Latin 2 groups are also best known by the Latin 4 groups. The opposite of this situation also obtains. The principles best known initially by all groups, and also best retained, are the use of the accusative case, place to which, and the predicate nominative. The principles least well known by all groups initially, and which also exhibit a definite tendency to be least well known at later test periods are: The dative of comparison, ablative with certain deponent verbs, partitive genitive, dative of possession, and the dative with certain intransitive verbs.

Conclusions

The more significant conclusions to be drawn from this investigation are:

1. The retention of identically correct responses over the summer vacation in the various groups ranged from six to fifty-one per cent in excess of the chance percentage. The stop groups retain identically correct responses over a period of one year from six to twenty per cent in excess of chance.

2. The remembrance of identical errors over the vacation periods exceeds the chance percentage from six to nineteen per cent in the various groups. The retention of identical errors over an interval of one year surpasses the chance percentage by from nine to twenty per cent. This suggests that some errors are well learned.

3. The range in retention of total correct responses over vacation is from sixty-six to eighty-five per cent in the given groups. Over a period of one year the range is from fifty-four to sixty-eight per cent.

4. Identical errors break down under the influence of study about three times more rapidly than identically correct responses are built up.

5. Time is more effective in destroying identically correct responses than identically incorrect responses.

6. The curves of retention, in general, correspond in form, but not in rate, to those of Ebbinghaus. From September to May 1930 the rate of forgetting is slow.

7. There is a suggestion that an intelligence quotient of 115 is the critical level in terms of mastery and retention of Latin syntax.

8. The analysis of the coefficients of correlation indicate that, in general, intelligence is not the determining factor in the acquisition of Latin syntax, and even in its remembrance intelligence is not the decisive element. Initial achievement is the significant factor in retention both from an absolute and relative viewpoint.

9. However, those who are in the upper quarter of intelligence know and retain more correct responses, both absolutely and relatively, than those who are in the lower quarter of intelligence.

10. Those who are in the upper quarter of intelligence retain a greater relative, but not absolute, amount of their initial errors.

11. Those who are in the lower quarter of intelligence retain, absolutely, more of their initial errors, but, relatively, fewer of them than those who are in the upper quarter of intelligence.

12. All reliability differences between the upper and lower quarters in terms of intelligence are statistically significant.

13. One month of study in the fall increases the mean of the Latin 2 continuation group 2.92 points beyond its pre-vacation mean. This is a statistically significant gain. The October mean score of the Latin 4 continuation group is 2.31 points below its pre-vacation mean, and only .46 points beyond its September mean. In no instance in this group is the reliability of the mean differences significant.

14. The Latin 2 non-modern language group is twelve per cent superior to the modern language group in retention of identically correct responses over the summer vacation.

15. Both the Latin 2 modern language and non-modern language groups retain total correct responses over the vacation period equally well, but in May 1930 the modern language group exceeds the non-language group by twenty-eight per cent. This difference is statistically reliable.

16. No significant differences were found in the Latin 4 group between the modern language and non-language groups.

17. Both the Latin 4 continuation and stop groups are significantly superior in all respects to both the Latin 2 groups.

18. No statistically reliable sex differences were found; however, the girls tend to excel the boys in all relations studied.

THE BOGEY OF THE "PERSONAL EQUATION" IN RATINGS OF INTELLIGENCE

HERBERT S. CONRAD

Institute of Child Welfare, University of California

It has become commonplace in psychology to suppose that subjective ratings of any trait are likely to be seriously affected by the "personal equation" of the judge. The theoretical possibility that the personal equation may render ratings incomparable is obvious enough; so obvious, in fact, that few (if any) writers have considered it necessary to bring forward rigorous proof that this theoretical possibility is a practical fact. A fair example of the attitude of most investigators may, we believe, be found in the monumental work of the Army psychologists, the *Memoirs* of the National Academy of Sciences, Vol. XV. In this volume, a discussion of the correlation between officers' ratings and intelligence test scores led to the following remarks:

> . . . if we place all individuals rated "6," for example, in the same class in our contingency table, *whether they are all rated by the same person or not*,[1] the computed measure of relationship (between ratings and test scores) will be affected by factors other than those relevant to the problem (*viz.*, by the "personal equation") . . . It is necessary, therefore, to equalize the mean ratings of all groups before they are combined. This may be done by treating the class intervals on the rating scale (of each judge) as intervals on the axis of abscissæ of a normal probability curve . . . "[2]

The Army psychologists adopted their own suggestion. Ratings by "at least five different rating officers"[3] on the men in five companies at Camp Meade, were adjusted on the basis of the normal curve,[4] and thrown into a single contingency- (or correlation-) table. The

[1] Italics ours.

[2] *Memoirs* of the National Academy of Sciences, 1921, Vol. XV, pp. 438–439. Throughout this article we shall adhere to the definition of the "personal equation" implied in this quotation from the *Memoirs; viz.*, spurious differences between judges in the means and standard deviations of their ratings.

[3] *Memoirs*, 439.

[4] The adjusted values of each point on the rating scale are given in a table on p. 441 of the *Memoirs*. Examination of this table shows that the "personal equation" in the five companies is, by ordinary standards, quite important, and ought to be minimized by some such method as that adopted by the Army psychologists.

correlation between officers' ratings and Army Alpha test scores on the total six hundred seventy-two men was then computed as .5867[1] (.6138 if corrected for limitation of range[2]). The same method was employed to determine the correlation between officers' ratings and Beta total score, on four hundred sixteen men in three companies at Camp Meade. The correlation between adjusted ratings and Beta score was found as .5798[1] (no correction for limitation of range was considered necessary in the case of the Beta test[3]).

Let us suppose, now, that no attempt whatever had been made to remove the personal equation: what would the correlation between ratings and test scores then be? *The correlations would be about the same.* Tables 110 and 118 of the *Memoirs* contain correlation-charts between ratings and test-scores for each company individually. Combining these individual correlation charts, and calculating the net correlation *without any adjustment whatsoever* of the ratings in each company, we obtain the following results (Table I):

TABLE I.—CORRELATION BETWEEN RATINGS AND TESTS, CAMP MEADE DATA

	n	Ratings unadjusted	Ratings adjusted to minimize the effect of the personal equation[1]
r, ratings × Alpha test.....	672	.5832 (.5938[2])	.5867 (.6138[2])
r, ratings × Beta test.....	416	.5953	.5798

[1] The figures in this column are taken from the *Memoirs*, p. 440.

[2] Corrected for the limited range of measuring-ability of the Alpha test—*cf.* *Memoirs*, pp. 439–440, 622–632.

It will be observed from Table I that correction of the ratings for the personal equation very slightly raises the correlation between ratings and Alpha; and slightly lowers that between ratings and Beta. The important fact, however, is the very slightness of these changes. Judging from the correlations between ratings and objective test scores, it is evident that *the attempted elimination of the personal equation has led to no appreciable improvement in the ratings.* The cause for this lack of improvement cannot be assumed to lie wholly, or even mainly, in any deficiency of the method of adjustment employed

[1] *Memoirs*, Table 101, p. 440.

[2] *Memoirs*, pp. 439–440, 622–632.

[3] *Memoirs*, p. 441.

by the Army psychologists.[1] The conclusion therefore seems inevitable that, at least in the present typical instance, *the personal equation effect (in terms of correlation) is relatively insignificant.*

These wholly unexpected findings in the Camp Meade material lead to a very important question: Has not the effect of the personal equation in ratings been grossly overestimated? Has not the misplaced emphasis upon the personal equation served as a kind of smoke screen, behind which the other, less easily located errors of ratings have escaped the careful scrutiny which they require? These questions appear especially significant, in view of the current widespread use of rating scales of all kinds for the measurement of attitudes and of character.

[1] The lowering of the correlation between ratings and Beta (Table I) suggests that in this instance the statistical conversion of the ratings in the various companies to the same normal curve—thus arbitrarily erasing whatever true differences among Companies which the ratings may have reflected—introduced an error greater than the one which the Army psychologists aimed to eliminate. Doubtless the statistical conversion should, in strict theory, have paid due regard to whatever true differences exist among the various Companies. As a matter of fact, however, the true differences among the Companies are really quite slight (*Memoirs*, Table 118). It follows that the error introduced by neglecting these true differences must have been small—and could exceed the effect of the personal equation, only because the personal equation effect is itself so very inconsiderable.

BOOK REVIEWS

The Influence of Habit on the Faculty of Thinking, by Maine de Biran, translated by Margaret D. Boehm. Baltimore: The Williams and Wilkins Company, 1929. Pp. 227.

English readers will appreciate the clarity with which the translator has made available this early nineteenth-century classic of French Psychology. It is the third in a series of translations and republications of important psychological works under the editorship of Knight Dunlap.

The introduction by George Boas sketches the development of the theory of habit in France and suggests the significance of Maine de Biran's essay in the movement from Condillac's seventeenth-century analytic sensation psychology to the twentieth-century spiritualism of Bergson.

Biran's essay on habit may be regarded historically as one contribution in the movement from a "passive" to an "active" psychology. Condillac, in his attempt to avoid superstitious overconfidence in the unexplained activity of a soul or mind, had built his theory upon sensations themselves. Although he assumed a soul, it was devoid of character (passive) until "modified" by sensations. While accepting Condillac's hypothesis, Biran was anxiously inquiring how far the soul becomes active and under what conditions. How can the repetition of sensations and movements conceived as habit lead to active reflective thinking? Coupled with this practical problem runs the question of the relationship between atomic sensations and the unity of the ego or soul. The problems attacked in the essay are fundamental; they still pursue psychologists in their efforts to reconcile habits and intelligence.

The essayist begins by distinguishing sensations which are passively felt from perceptions in which active seeking movements occur. He contrasts seeing with looking, hearing with listening. Out of the effort involved in movements comes a recognition of the ego, and thereby a separation of the "I" from the "not I," which makes knowing possible. Repetition of movements (habit) not only brings forth the ego, but gives it stability by constituting a persisting memory. With this introduction Biran is ready to delve into the complex effects of habitual action in two sections entitled respectively "Passive Habits" and "Active Habits."

Biran explains how sensory impressions grow weaker when they are continued or frequently repeated, because the equilibrium of the organism is less and less disturbed. For the organism tends to restore its equilibrium either by lowering the natural tone of the stimulated organ or by raising that of the unstimulated parts. Thus through habituation the effect of particular stimulations permeates the whole organism, but the specific impression seems to grow constantly weaker. Likewise, the effort felt in movement becomes less as these movements become habitual, although at the same time these habits are becoming more easy, prompt, and precise. So finally a movement "manifests itself no longer to consciousness except by the results" produced. "It is thus . . . that habit effaces the line of demarcation between voluntary and involuntary acts." Nevertheless, the brain as a central organ subsumes the partial activity of the organs and the habits characteristic of them in one framework, thereby promoting an adjustive precision in action. The impressions received by sight and touch, for example, are "confused by habit in an indivisible perception." The author makes clear the two contrasting aspects of habit—"insensibility in action" and "precision in results"—out of which comes the practical problem whether habit shall be a base of experience which directs us or "a too fertile source of prejudice which blinds us."

Under "Active Habits" Biran discusses the symbols employed in memory and the effects of repetition thereon. He distinguishes three kinds of memory. *Mechanical* memory develops when the methods of teaching draw attention too exclusively to the symbol so that the meaning disappears. In *sensory* memory the symbol attains an excitive power which produces emotional feelings especially through metaphors, allegories, and other immoderate figures. The third type, *representative* memory, gives symbols the infallible power of evoking ideas or images of distinct perceptions with fidelity. "This is the only useful memory and the sole basis of human intelligence." The escape from automatism is finally made by stating that habit gives thought "activity to apprehend and appropriate all that conforms to it." This conformity may occur through analogy, which "makes the transition from the known to the unknown so gentle." "We are in a new world, and it seems that we have not left the precincts of our habits."

The student of EDUCATIONAL PSYCHOLOGY, whether he is concerned chiefly with the current interpretations of habit or with the historical

development of the concept, will find an acute challenge to his discriminatory powers in the essay of the young Maine de Biran. (He was thirty-three when the essay was first offered in the 1799 prize competition of the French Institute.) The reader must judge how much Biran overloads the habit concept. Critical reading of this classic will bring a deeper appreciation of the complex problems which are too often concealed within the term habit. WILLIAM F. BRUCE.

State Normal School, Oneonta, N. Y.

Child Psychology, by Margaret Wooster Curti. New York: Longmans, Green and Company, 1930. Pp. IX + 527.

That child study continues to develop as the science of human behavior is evidenced by Dr. Curti's comprehensive and scholarly treatment of the subject. Few writers have succeeded so well in combining the theoretical and practical aspects of the subject: Too often there is plenty of "arm chair" thinking, but little practical application; or on the contrary, a wealth of anecdote, case study material and illustration with a paucity of organized generalization. Dr. Curti happily strikes the golden mean, producing a volume that is both readable and authentic. Topics too often slighted are given ample treatment. This is true particularly of such topics as the child's inheritance, the relation of physical and mental growth, the causes and control of antisocial conduct, and the development of personality. Other topics included are the more familiar subjects of instincts, learning, reasoning, thinking and play.

The author treats each topic with originality and distinction, professing allegience to no one school of psychology, but selecting from each the contributions of greatest scientific value. Her expressed aim is to survey the facts of child behavior in order to provide a rational basis for child guidance. Older theories are reviewed in the presentation of each topic, along with the newer results of research, but no theories are accepted by the author as tenable when they are not in harmony with established fact. The author prefers the term "native responses" to "instincts," of which she lists several groups or classes. Interesting first-hand accounts of conditions reflex experiments with children are described. With other psychologists, the author finds the conditioned reflex principle too limited to explain the whole of learning, though she considers it a useful concept for understanding and explaining certain phases of the learning process.

The sections on maladjustment and delinquency show keen insight into child adjustment problems. The author is inclined to attribute most antisocial behavior to thwarting of desires and unfavorable environmental conditions. At this point she verges closest on dogmatic attitude, for she seems disinclined to accept any theory not in harmony with these principles and is even skeptical of facts which appear to indicate that heredity, defective mentality, and other factors may be the causal ones. In her opinion punishment should not be retaliative nor proportionate to the enormity of the offense, but should be in harmony with the causes back of the offense. A list of selected references for each chapter is given at the end of the book.

As a basic text in child psychology the book can be well recommended. Most profit will be derived from it by those who have thorough grounding in the principles of psychology, who are familiar with the terminology and divergent theories of different schools, but to some extent the book itself furnishes this background.

GERTRUDE HILDRETH.
Lincoln School of Teachers College, Columbia University.

Objectives and Procedures in Civic Education, by Charles Clinton Peters. New York: Longmans-Green and Co., 1930. Pp. X + 302.

Of the listing of objectives there seems to be no end, but as Peters states in his preface, we have hardly made a beginning of doing anything about the objectives after they were stated. The transition from objectives to teaching programs is often made very superficially, if at all.

This volume attempts to do something very definite in the way of getting civic objectives into practice as a basis for actual civic education. In addition to this aim, there is another equally important, namely to illustrate the method of attack upon curriculum problems, that may equally well be made in connection with other divisions of the curriculum. In other words, Peters has selected a very important branch of the curriculum upon which to do a good job, and offers that job as a model or pattern for further research and practical curriculum making.

There are chapters on the purposes and social objectives of education for citizenship, psychological objectives as a means of civic education in the school, and the contributions to civic education

from the different subjects of the curriculum. In addition there are suggestions for education for citizenship at various grade levels and for adults. The final chapter, and also a very interesting one, deals with education for industrial and social democracy.

The basic procedure for the entire work is what the author calls the making of a blue-print of an optimum citizen. This consisted of having a number of graduate students and other thoughtful adults each get in mind a particular person whom he considered to be an ideal citizen, and to list the qualities or actions of that person which distinguished him from less desirable citizens. Likewise, particular individuals representative of poor citizenship were described for their faults or deficiencies. In this way a very complete list was made, and additional persons were included to the point of diminishing returns, so that practically no new items were added for quite a while before the listing was discontinued. In this way the composite list is assured of practical completeness.

The qualities of the good citizen are listed as points in an outline, and after each point are listed possible means and occasions for training in these abilities. In a third column the author gives comments on methods and procedure for giving this training.

The author's actual blue-print of the good citizen occupies only forty pages of the book (pp. 56–97). The remainder of the volume is devoted to practical discussions of ways and means whereby this ideal citizen may be developed.

Toward the end of the volume there is a similar outline which the writer calls a blue-print of social democracy, and which the student who uses the book is expected to work out in the three-column form to show how each of the qualities of the ideal state may be developed.

The work is rich in concrete and helpful material, the general idea behind it is good, and the curriculum-making technique exemplified has obvious usefulness in other subjects. C. C. CRAWFORD.

University of Southern California.

Present Day Psychology: An Objective Study in Educational Psychology, by Alene Ralston and Catharine J. Gage. Philadelphia: J. B. Lippincott Co., 1931. Pp. XIV + 404.

This book is intended for the uninitiated, the parent in the home, and the teacher in elementary school service. It consists of three

parts—Living, Learning, and Measuring. The authors state that they have maintained an objective view point and that no statement has been made unless it has been verified by authorities through experimentation. If this be so, then they have extended the meaning of experimentation; for ideas, concepts, beliefs, libido, œdipus complex, love (of self, parents, comrades, and a mate), attitudes, sentiments, hypnotism, hysteria, introversion, extroversion, egoism, etc., all find a place. A very large number of authors and studies are quoted and this is at once the strength and weakness of the volume—its strength in that a very catholic point of view has been maintained; its weakness in that the uninitiated will tend to be confused by the multiplicity of outlooks. The section on the Mendelian Law has confused average results with specific ones (see for example, black mice on page 222) and there are other mis-statements of fact in this section. On the whole, however, the authors have done their work well. Persons skilled in the subject will find much to interest them and many novel treatments and suggestions. The book is well bound, well printed, and well documented.

PETER SANDIFORD.
University of Toronto.

Mental Aspects of Stammering, by C. S. Bluemel. Baltimore: The Williams and Wilkins Company, 1930. Pp. X + 152.

Stammering (stuttering) is considered in this book as "an impediment of thought and not as an impediment of speech." In a measure, the discussions supplement the point of view expressed in the author's previously published work, "Stammering and Cognate Defects of Speech," but is written primarily for school use and as a manual for the stammerer who must rely on self training, consequently the greater part of this volume is devoted to speech drills and exercises for "training thought as the basis of speech correction." Part I contains a brief exposition of the author's theories concerning the nature and origin of this speech disorder.

Heretofore Dr. Bluemel has considered the stammerer's difficulty as an inability to speak without hesitation occasioned by a momentary "amnesia" or loss of memory for the word he wishes to say. Now he modifies this conception by the notion that the difficulty is not so much an inability to *recall* the sound as it is a *recoil* of the sound from the mind. Amnesia occurs, not because the stammerer waits upon his memory, but because he is confronted with a break in the

continuity of consciousness. The characteristic spasms of stammering are to him a symptom rather than a cause of the impediment and are "nothing more than the abortive struggle that results from the stammerer's efforts to speak with unaided motor images."

Recommendations for therapeutic treatment are influenced by the idea that "underlying all this disturbance of thought there is, of course, a neurotic temperament that makes recoil possible and that the larger task of the stammerer is not merely to tranquilize his mental speech but to tranquilize his mind and his life." This tranquilization comes through "mental drill" in which "thoughts are disciplined like soldiers," by being brought first into line, then into step, and finally into orderly manoeuver. Thereby does the stammerer learn to focus his attention more sharply on the mental process of speech, thus avoiding recoil of imagery and incoordination of thought. Drills are composed of attempts to speak to signals, speaking in unison, "silent" or "phantom speech," "relaxation," "slow speech," and breathing exercises. The tasks are graded for difficulty and situations requiring speech are merged from the less difficult to the more exacting.

Agreement with Dr. Bluemel's theory presupposes an acceptance of the notions that physical and mental aspects of speech are disparate or dual in nature rather than two phases of one function; that individuals have a predominant type of imagery (visual, auditory, or kinesthetic) in "the stream of consciousness," and that there is a neurotic temperament present in the stammerer. Evidence contradictory to these ideas has appeared in sufficient amounts to make the hypothesis of "auditory amnesia and recoil" somewhat untrustworthy.

At any rate Part II of this volume contains a description of speech exercises which have for many years proved their usefulness in improving articulation and enunciation of language whatever may be the causes of the speech difficulties being eradicated. The system of signals is probably the only one of questionable merit for use in the classroom. Consequently the book will be a valuable addition to the library of a teacher.

ELIZABETH D. MCDOWELL.
Teachers College, Columbia University.

The Psychology of the Common Branches, by William Henry Pyle. Baltimore: Warwick and York, Inc., 1930. Pp. VIII + 381.

There are several unique features about this book. It is primarily a collection of abstracts of source material, but the job has been done

in a rather unique fashion. Each chapter has four main divisions, as follows: (1) The text proper, in which are stated the main facts with such discussion as seems necessary; (2) exercises and problems for the students; (3) concise, numbered statements of all the essential principles; (4) abstracts of the articles and books in the field. The first section resembles the usual chapter in the ordinary book, being the author's own direct and very effective discussion of the topic of the chapter. The second division, namely the exercises, consists of laboratory experiments for the student to do and thought questions for him to answer. These are very ingeniously prepared. The third section is usually about a page or two in length, and contains twenty or thirty concise sentences which sum up the main points of all the experimental work available on the subject. The last section in each chapter is a bibliography briefly annotated to show the contribution of each writer to the general problem of the chapter.

This style of treatment makes the volume much more useful for teaching purposes than it would have been if the author had tried to keep each individual unit of source material intact. The reader can get the important points from all the source material without having to digest a number of overlapping and contradictory investigations. In other words, the author has digested and blended this miscellaneous mass of incoherent matter, has omitted duplications, and has ironed out conflicts and inconsistencies very skillfully.

The few samples of the author's excellent summary of essential principles, taken from the handwriting chapter, will illustrate how well he has formulated the results of objective investigations into useful principles. We quote from the portion dealing with the physiology and hygiene of handwriting:

6. Teaching writing involves not only the movement of making letters, but posture, position of paper, manner of holding the pen, and the tools of writing, paper, pen, ink.
7. Learning to write is the process of acquiring a large number of motor habits.
8. The paper should slope to the left about 45°.
9. The light should be diffused daylight.
10. The pupil should face the desk squarely.
11. In good posture the body and head are held erect.
12. Good posture in writing is a habit which must be built up.
13. Almost the whole of both fore-arms should rest on the table.
14. The first pen used should be coarse, and the holder made of material easily held, as cork or rubber.

Some examples of the exercises for the student, taken from the chapter on reading, are as follows:

1. Determine how much vocalization you have in suppressed form in your silent reading.
2. See if by practice you can decrease vocalization and increase the speed of silent reading.
3. Try your speed in silent reading in reading many different types of material.
4. Try changing your attitude or purpose in reading the same kind of material. What do you discover?
5. Make a complete list of devices for arousing interest in reading.
6. Try to discover the source of your pleasure in several different pieces of literature, poems, essays, fiction, etc.
7. Make a study of word meanings, with the object of discovering what constitutes meanings.

One serious limitation should be mentioned. The book treats only four of the elementary subjects, reading, spelling, handwriting, and arithmetic. The splendid service rendered to these four subjects might well be rendered also with respect to language, drawing, music, handwork, geography, citizenship, and others. Perhaps the author will do this in a later volume.

C. C. CRAWFORD.
University of Southern California.

PUBLICATIONS RECEIVED

ANDERSON, LEWIS FLINT: *Pestalozzi.* New York: McGraw-Hill Book Co., Inc., 1931, pp. IX + 283.

AVELING, FRANCIS: *Personality and Will.* New York: D. Appleton and Co., 1931, pp. X + 246.

BACHMAN, FRANK P., Editor: *The All-year School of Nashville, Tennessee.* Nashville: George Peabody College for Teachers, 1931, pp. 60 (paper).

BADANES, SAUL: *A Child's Third Number Book.* Part Two. New York: The Macmillan Co., 1931, pp. 141.

BARR, A. S. and MABEL RUDISILL: *An Annotated Bibliography of the Methodology of Scientific Research as Applied to Education.* Madison: University of Wisconsin, 1931, pp. 129 (paper).

BASSETT, CLARK: *The School and Mental Health.* New York: The Commonwealth Fund, 1931, pp. 66 (paper).

Behavior Problems of School Children. New York: The National Committee for Mental Hygiene, Inc., 1931, pp. 20 (paper).

BOLDYREFF, TATIANA W.: *By Word of Mouth.* Boston: The Gorham Press, 1931, pp. 144.

BRUBACHER, JOHN S.: *Henry Barnard on Education.* New York: McGraw-Hill Book Co., Inc., 1931, pp. X + 298.

CARR, WILLIAM G. and JOHN WAAGE: *The Lesson Assignment.* Stanford University, Calif.: Stanford University Press, 1931, pp. VII + 98.

COOK, WALTER WELLMAN: *The Measurement of General Spelling Ability Involving Controlled Comparisons between Techniques.* Iowa City, Iowa: University of Iowa, pp. 112 (paper).

Education and Racial Adjustment. Report of Peabody Conference on Dual Education in the South. Atlanta, Ga.: Executive Committee of the Conference, Standard Building, 1931, pp. 78 (paper).

ENG, HELGA: *The Psychology of Children's Drawings.* New York: Harcourt, Brace and Co., 1931, pp. 223.

EURICH, ALVIN C.: *The Reading Abilities of College Students.* Minneapolis: The University of Minnesota Press, 1931, pp. 208.

Examining the Examination in English. Cambridge: Harvard University Press, 1931, pp. XV + 295.

FOX, CHARLES: *The Mind and Its Body.* New York: Harcourt, Brace and Co., 1932, pp. 316.

FRYER, DOUGLAS: *The Measurements of Interests.* New York: Henry Holt and Co., 1931, pp. XXXVI + 488.

HARRIS, DANIEL: *The Relation to College Grades of Some Factors Other Than Intelligence.* New York, 1931, pp. 55 (paper).

KARPF, MAURICE J.: *The Scientific Basis of Social Work: A Study in Family Case Work.* New York: Columbia University Press, 1931, pp. XX + 424.

KAWIN, ETHEL, and CAROLYN HOEFER: *A Comparative Study of A Nursery-school versus A Non-nursery-school Group.* Chicago: University of Chicago Press, 1931, pp. 52 (paper).

KEATINGE, M. W.: *Comenius.* New York: McGraw-Hill Book Co., 1931, pp. 255.

LOWREY, LAWSON G., Editor: *Institute for Child Guidance Studies.* New York: The Commonwealth Fund, 1931, pp. VIII + 290.

LUMLEY, FREDERICK E. and BOYD H. BODE: *Ourselves and the World.* New York: McGraw-Hill Book Co., Inc., 1911, pp. VIII + 591.

MACIVER, R. M.: *Society—Its Structure and Changes.* New York: Ray Long and Richard R. Smith, Inc., 1931, pp. XVI + 569.

MAGENIS, ALICE and MADELINE F. GILMOUR: *Directed High School History Study.* Yonkers, N. Y.: World Book Co., 1932, pp. 188 (paper).

MARSHALL, M. V.: *Education as A Social Force.* Cambridge: Harvard University Press, 1931, pp. XXIX + 161 (paper).

Mental Hygiene in the Classroom. New York: The National Committee for Mental Hygiene, Inc., 1931, pp. 44 (paper).

MOORE, BRUCE V. and GEORGE W. HARTMANN, Editors: *Readings in Industrial Psychology.* New York: D. Appleton and Co., 1931, pp. XXXIX + 560.

MOORE, ELIZABETH SKELDING: *The Development of Mental Health in a Group of Young Children.* Iowa City, Iowa: University of Iowa, 1931, pp. 128 (paper).

O'CONNOR, E. A.: *Local School Problems.* Little Falls, N. Y.: The Author, 1931, pp. 24 (paper).

PRESCOTT, GERALD W.: *Iowa Algae.* Iowa City: University of Iowa, 1931, pp. 235 (paper).

PRESSEY, LUELLA COLE and JESSIE MARY FERGUSON: *Student's Guide to Efficient Study.* New York: Richard R. Smith, Inc., 1931, pp. 39 (paper).

RASMUSSEN, VILHELM. *Diary of a Child's Life.* London: Glydendal, Ltd., pp. 187.

ROBERTSON, J. M.: *A History of Freethought.* New York: G. P. Putnam's Sons, 1930, Vol. I, pp. XXXV + 312. Vol. II, pp. 322.

SMITHEY, W. R., Editor: *Secondary Education in Virginia.* Charlottesville: University of Virginia, 1931, pp. 71 (paper).

STETSON, F. L.: *The Organization and Administration of Standard High Schools.* Eugene: University of Oregon Press, 1931, pp. 205–296 (paper).

STROMSTEN, FRANK A.: *The Development of the Gonads in the Gold Fish, Carassius Auratus (L.).* Iowa City: University of Iowa, 1931, pp. 45 (paper).

TRAVIS, LEE EDWARD: *Speech Pathology.* New York: D. Appleton and Co., 1931, pp. XXXIV + 331.

WOLF, MARGARET M.: *The Polyporaceae of Iowa.* Iowa City: University of Iowa, 1931, pp. 93 (paper).

YEAXLEE, BASIL A.: *The Approach to Religious Education.* New York: The Macmillan Co., 1932, pp. 144.